Signaling, Screening and Self-Selection in Student Scholarship Applications – An Empirical Investigation of Stipend Awarding Decisions in Germany –

Der Fakultät für Wirtschaftswissenschaften der
Universität Paderborn
zur Erlangung des akademischen Grades
Doktor der Wirtschaftswissenschaften
- Doctor rerum politicarum -
vorgelegte Dissertation
von

Linda Sophia Kurze

geboren am 27.06.1985 in Lübeck

2013

VORWORT

An dieser Stelle möchte ich meinen Dank aussprechen an all diejenigen, die mich während der Bearbeitung der Dissertation begleitet haben und mir Mut, Kraft und Zuversicht gaben, um den eingeschlagenen Weg zu gehen und konsequent weiter zu verfolgen.

Meinem Doktorvater Prof. Dr. Bernd Frick für die optimale Betreuung, sein Vertrauen und die produktive Zusammenarbeit. Er gab mir immer die Möglichkeit eigene Ideen umzusetzen und half mir durch inspirierende und konstruktive Kritik diese umzusetzen. Prof. Dr. Martin Schneider, Prof. Dr. Esther Winther und Prof. Dr. Burkhard Hehenkamp für ihre Bereitschaft als Zweitgutachter bzw. Promotionskommissionsmitglieder zu fungieren und für die fachlich äußerst bereichernde Diskussion, nicht nur im Rahmen der Disputation.

Meinen kompetenten, hilfsbereiten und interessierten Ansprechpartnerinnen beim Deutschen Akademischen Austauschdienst, Dr. Simone Burkhart und Gabriele Knieps.

Meinen lieben ehemaligen und aktuellen Kollegen Dr. Christian Deutscher, Dr. Filiz Şen-Polat, Dr. Marcel Battré, Arne Büschemann, Andre Kolle, Friedrich Scheel und Anica Rose für eine Zusammenarbeit, die weit über die berufliche Ebene hinaus geht, sowie für ein freundschaftliches und humorvolles Arbeitsklima, das sowohl zu Beginn auf dem C-Flur der Universität sowie die letzten drei Jahre im Nordflügel des Technologiparks 21 die Strapazen des Promovierens immer wieder versüßt hat.

Den studentischen Hilfskräften danke ich für ihre tatkräftige Hilfe bei jeglichen unterstützenden Tätigkeiten im Zusammenhang mit meinem Datensatz. Einer von euch muss ich an dieser Stelle aber den größten Dank aussprechen: Anja Viehmann, herzlichen Dank, dass du mich so tatkräftig bei der Dateneingabe vor Ort in Bonn unterstützt hast. Ich bin froh, dass die vier intensiven Wochen der „Datenklopperei“ unserer Freundschaft keinen Abbruch getan haben.

Meinen Eltern, meiner Familie und meinen Freunden, sowie meinem Lebensgefährten Marco möchte ich für ihre Liebe und Unterstützung meinen ganz besonderen Dank aussprechen. Sie haben mich stets motiviert und an mich geglaubt und mir dadurch erst einen solchen Erfolg ermöglicht.

Paderborn, im Februar 2013 Linda Kurze

TABLE OF CONTENTS

List of Abbreviations

AGG	Allgemeines Gleichbehandlungsgesetz
AIC	Akaike Information Criterion
BAföG	Bundesausbildungsförderungs-Gesetz
BMBF	Bundesministerium für Bildung und Forschung
CAD	Canadian Dollars
CDF	Cumulative Distribution Function
CEO	Chief Executive Officer
CHE	Centrum für Hochschulentwicklung
CPA	Certified Public Accountant
CRQ	Centre for Research into Quality
CV	Curriculum Vitae
DAAD	Deutscher Akademischer Austauschdienst
DAS	Datenarchiv für Sozialwissenschaften
Destatis	Statistisches Bundesamt
DFG	Deutsche Forschungsgemeinschaft
DIW	Deutsches Institut für Wirtschaftsforschung
ECA(s)	Extracurricular Activity (-ies)
Ed.	Edition
EEOC	Equal Employment Opportunity Commission
FFM	Five-Factor Model
GESIS	Gesellschaft für Informationssysteme mbH
GoF	Goodness of Fit
GPA	Grade Point Average
GRE	Graduate Record Examination
HIS	Hochschul-Informations-System mbH
HR(M)	Human Resources (Management)
IFS	Institut für Schulentwicklungsforschung
ILO	International Labour Organization
IM	Impression Management
ISBM	Institute for the Study of Business Markets
I-SE	Interviewing Self-Efficacy
IZA	Institut zur Zukunft der Arbeit
KH	Katholische Hochschule
KMK	Kultusministerkonferenz
KSA	Knowledge(s), Skill(s) and Ability (-ies)

LOR(s)	Letter(s) of Reference/Recommendation
LPM	Linear Probability Model
LR	Likelihood-Ratio
MBA	Master of Business Administration
NAACP	National Association for the Advancement of Colored People
NBER	The National Bureau of Economic Research
NEO PI	Neuroticism-Extraversion-Openness Personality Inventory
NFER	National Foundation for Educational Research
NSF	National Science Foundation
OECD	Organisation for Economic Co-Operation and Development
OLM	Ordered Logit Model
OLS	Ordinary Least Squares
P-E	Person-Environment
PH	Pädagogische Hochschule
P-J	Person-Job
PNAS	Proceedings of the National Academy of Sciences
P-O	Person-Organization
P-P	Person-Person
pp(s)	Percentage point(s)
QCA	Qualitative Comparative Analysis
RD	Regression Discontinuity
ML	Maximum Likelihood
RIASEC	Realistic-Investigative-Artistic-Social-Enterprizing-Conventional
S & E	Science and Engineering
SAT	Standard Aptitude Testing
SHRM	Society for Human Resource Management
SULCIS	Stockholm University Linnaeus Center for Integration Studies
TH	Technische Hochschule
THE	Times Higher Education
TOEFL	Test of English as a Foreign Language
UCAS	Universities and Colleges Admissions Service
UCI	University of California Irvine
UK	United Kingdom
US	United States of America
VGRdL	Volkswirtschaftliche Gesamtrechnung der Länder
VS	Verlag für Sozialwissenschaften

LIST OF TABLES

LIST OF FIGURES

PART A: RESEARCH QUESTION AND RELEVANT LITERATURE

1 INTRODUCTION

In Germany, the shortage of skilled labor is currently one of the most important topics of conversation in both politics and the corporate sector (Brenke 2010). Both enterprises and employers' associations caution against the increasing long-term lack of highly skilled workforce. The anticipated shortage will mainly be caused by the demographic change as well as the increasing emigration of highly educated citizens (Zimmermann 2010). Although empirical studies conducted by e.g. the DIW ('Deutsches Institut für Wirtschaftsforschung') and the Federal Employment Agency (Brenke 2010; Bundesagentur für Arbeit 2011) do not confirm an immediate short-term shortage, the Federal Government tries to counteract this trend as early as possible. By setting the right incentives, the Federal Ministry of Education and Research (BMBF) tries to increase the number of well-trained citizens in Germany. One possibility to generate a more highly skilled workforce is to encourage young citizens to upgrade their education by attending a university. In order to convince more (promising) high school graduates to enroll in tertiary education, the BMBF has increased public spending on merit-based scholarships, i.e. scholarships that are allocated on the basis of individual achievement. In 2009, 132.3 million € were available for merit-based student scholarships (BMBF 2009). In comparison to 2005, this corresponds to a 64% increase in financial means dedicated to merit-based stipends.

Albeit, only very little is known about the selection criteria applied in merit-based scholarship. Who among all German students is awarded a scholarship and why? The BMBF usually only claims to select the 'best and most promising' students based on elaborated selection processes, but virtually nothing is known about the selection criteria employed in this process. The BMBF (2009) merely states to select on the basis of subject-specific achievement, personality as well as extracurricular activities. How applicants are able to signal their ability and how evaluators screen potential scholarship holders is basically unknown. In order to be able to assess the effectiveness of these measures, however, one needs to gain in-depth knowledge of the selection process and its outcome. Only then, it is possible to evaluate whether or not merit-based financial aid is indeed capable of acting as a remedy for the imminent shortage of skilled labor by supporting promising young executives and elite members.

The only previous attempt to shed some light on the characteristics of German merit-based scholarship holders has been conducted by Middendorff, Isserstedt and Kandulla (2009). Using survey data collected from current stipend awardees they identified how the 'average' German scholarship holder looks like. Although they were able to get in-depth insight into the characteristics of those scholarship holders that answered to their request (approximately 50% of all contacted subjects), they were not able to separate self-selection (supply-side) from screening (demand-side) effects. In order to be able to differentiate between supply- and demand-side effects, the present research adopts a different approach and analyzes empirically how stipend awarding in one single scholarship granting organization is conducted. Being aware of the distinct selection criteria used in stipend awarding, the reader will then have a detailed idea about the pool of stipend awardees which – according to the BMBF's vision – represents the group of future executives and elite members.

In order to be able to answer the prevalent research question, this thesis is structured as follows: In **Part A** (Research Questions and Relevant Literature; Chapters 1-3), the reader will not only become aquainted with the main research questions as well as the institutional background of stipend awarding decisions, but also relevant literature will be presented. **Part B** (The Determinants of Successful Scholarship Applications: Theoretical Considerations; Chapters 4 and 5) subsequently provides a theoretical basis for an understanding of the drivers behind stipend awarding decisions. Finally, the hypotheses derived from Part B will be tested empirically in **Part C** (Empirical Evidence on the Determinants of Successful Scholarship Applications; Chapters 6-8). In more detail, **Chapter 2** provides a basis for understanding stipend selection processes by defining the different types of financial aid available to German students. **Chapter 3** then summarizes the available previous literature. Here, not only previous research on stipend awarding decisions will be presented, but also success factors of other types of selection processes will be discussed. Subsequently, the theoretical background of the investigated selection decision will be presented in **Chapter 4**. In this chapter, the decision of whom to award a stipend will be discussed both from a sociological and an economic perspective. Based on what has been addressed in Chapter 3 and 4, a conceptual model reflecting the current selection situation will be developed and a number of testable hypotheses will be derived in **Chapter 5**. The empirical part (Part C) begins with the presentation of the data set which has been specifically compiled for this research project (**Chapter 6**). **Chapter 7** then addresses the actual empirical analysis of stipend awarding success whereas **Chapter 8** concludes by providing a brief summary of the main results as well as addressing limitations and future research possibilities.

2 FUNDAMENTALS OF STUDENT STIPEND AWARDING

Before analyzing distinct success factors of scholarship selection processes, the reader needs to become acquainted with the concept of higher education stipends.[1] Consequently, the present chapter addresses the essentials of student scholarships by first providing a definition as well as discussing several types of financial aid. Subsequently, some peculiarities of stipend awarding in Germany will be presented.

2.1 DEFINITIONS AND TYPES OF STUDENT STIPENDS

Whenever an individual desires to visit a tertiary education institution, he or she needs to ensure sufficient funding for several years. Students – or in most cases their parents – typically do not only need to pay tuition and general fees, but also have to cover their day-to-day expenses for a period of three to five years. Several possibilities of financing higher education exist: Working either full- or part-time alongside their studies or raising a loan are only two of the opportunities students have. On the other hand, student stipends have also been increasingly offered by several institutions and private organizations. A stipend in general can be defined as a "source of funds that is provided to an individual, […] which allows the individual to pursue a particular interest" (Business Dictionary 2012). In the present case, the 'particular interest' would be to study or more precisely to earn a higher educational degree by graduating from university. Another (more functional) definition describes a stipend as a "predetermined amount of money that is provided periodically to help offset expenses" (Investopedia 2012). Depending on the scholarship provider, college scholarship programs can either be state-funded or privately-organized, but only state-based financial aid programs will be discussed here.[2]

The policy goal of each publicly funded scholarship program is usually to enable more high school graduates to study at a higher education institution, i.e. "to make the opportunity for a college enrollment more accessible" (Duffourc 2006, 236). Hence, an increase in college enrollment is usually aspired. Several distinct types of stipends have emerged from this initial objective: By reference to the basis on which scholarships are awarded one needs to differentiate between need-based and merit-based scholarships.

---

[1] In the course of this thesis, the terms 'stipend', 'scholarship', 'sponsorship', 'grant' and 'award' will be used interchangeably although slightly different definitions exist in the literature.

[2] As selection in privately funded programs might be bound to and hence awarded on the basis of the idiosyncratic ideas of the stipend provider, selection and retention criteria are very likely to be biased by the private organization's vision.

Need-based – also called means-tested scholarships – are allocated on the basis of somebody's individual financial neediness (Duffourc 2006; Federal Student Aid). Based on the student's – or in most cases his or her parents' – discretionary income, need-based financial aid is provided as a grant for low-income students only. Consequently, the student's ability to pay determines the amount of financial assistance (Monks 2009). Apparently, the objective of these programs is to facilitate university or college attendance for otherwise under-privileged high school graduates (Duffourc 2006; Monks 2009). Merit-based scholarships in contrast are awarded on the basis of individual (previous) achievement such as academic, artistic, athletic or other abilities. More specifically, these programs can be understood as a "form of financial assistance that does not require repayment or employment and which is usually offered to students who show potential for distinction, or who possess certain characteristics important to the scholarship provider (such as religious beliefs, hobbies, ethnicity, etc.)" (Peterson 2008[3]). Selection for (state-funded) merit-based programs in the US is mainly executed on the basis of standardized academic qualifications such as GPA or SAT scores (Duffourc 2006). It becomes obvious that merit-based financial aid does not pursue the objective of simply raising college enrollment per se, but to attract "worthy high school graduates" (Duffourc 2006, 235), i.e. "the most academically desirable applicants" (Monks 2009, 99). The policy goal of each federal state therefore is to keep the best, i.e. the "academically proficient students" (Cornwell, Mustard and Sridhar 2006, 761), in state.

Both types of state-based financial aid usually compete for the same funds and are consequently permanently criticized by the respective opponents. Means-tested programs are entirely oriented toward needy families. As a result, children whose parents' income is (slightly) above the threshold level are excluded by definition. Due to the positive correlation between socio-economic status and high school grades, merit-based scholarships however are criticized for disproportionally supporting "already-advantaged students" (Duffourc 2006, 244). As a result of the prevailing need-merit debate, most countries opt for a combination of both types of student financial aid. In addition to the need-merit distinction, scholarships can also be classified as e.g. student-, career- or college-specific according to the person, institution or purpose they are bound to. In contrast to the United States, German higher education is mainly taxpayer-funded and little, if any tuition is charged at public institutions. Nevertheless, both need- and merit-based student scholarship programs exist which will be briefly presented in the following section.

[3] http://www.fastweb.com/financial-aid/articles/815-financial-aid-glossary?page=5.

2.2 Stipend Awarding in Germany – Facts and Figures

Owing to the vision that higher education should be accessible to everybody, most publicly-funded German universities do not charge tuition fees, but are funded with the help of public means only, i.e. they are indirectly financed by each taxpayer.[4] Still, students are in need of funds in order to be able to pay their everyday expenses.

In order to enable as many high school graduates as possible to enroll in tertiary education, a nation-wide means-tested program has been established. Each student whose parents' income and/or financial assets are below a certain threshold is entitled to benefit from the student grant and will be allocated the grant upon application (BMBF 2010). The program is regulated by a law named 'Bundesausbildungsförderungs-Gesetz' (BAföG) and this acronym is also used in order to address the financial aid per se. The German legislation has decided upon certain amounts the 'typical' German student needs in order to cover living as well as study-related expenses. This amount is adapted on a regular basis. At present, the maximum rate for students who do not live with their parents adds up to 670€ per month (www.bafoeg.bmbf.de). Dependent on the individual financial background, each student's financing gap is calculated by subtracting this student's disposable income from the maximum rate. 50% of the financial aid provided by the BAföG is designed as a grant, i.e. this part is not repayable by the student (BMBF 2010). The other half of the financial aid is designed as an interest-free loan. This loan needs to be repaid after graduation, but only up to the maximum amount of 10,000€ (every amount which has been disbursed over and above these 10,000€ is transformed into a non-repayable grant). Redemption is not possible until five years after graduation and might be paid by installments of at least 105€ per month (BMBF 2010). Students however whose financial assets exceed the threshold level are not entitled to be sponsored by the BAföG. Hence, only 'under-privileged' students benefit from the aid regulated by the BAföG. In 2007, almost 500,000 students were sponsored by the BAföG program (Schmidt 2009). This corresponds to 25.5% of all students enrolled at any type of higher education institution in Germany (Destatis 2008; Schmidt 2009).

In order to additionally support particularly talented students regardless of their socio-economic background, several state-funded organizations also award merit-based scholarships to German students. In particular, twelve organizations who provide scholarships for outstanding students ('Begabtenförderungswerke') allocate merit-based stipends with funds provided by the Federal Ministry of Education and Research.

[4] In some federal states tuition fees are charged. Albeit, they do not exceed 500€ per semester.

Each of these twelve organizations[5] puts emphasis on different applicant characteristics (BMBF 2009, 2011) and pursues slightly different goals (reflecting the pluralistic concept of German merit-based sponsorship). The 'Cusanuswerk' as well as the 'Evangelische Studienwerk e.V. Villigst' for instance are church-related institutions and consequently aim at sponsoring catholic (or protestant) students. Other organizations, e.g. the 'Friedrich-Ebert-Stiftung', are rather affiliated to distinct political parties and desire to support prospective political leaders (BMBF 2009, 2011).

In order to further increase the number of stipend awardees, the BMBF started an additional stipend program in 2011 called 'Deutschlandstipendium'. This type of merit-based stipend is especially directed at elementary students promising to perform excellently at university and during their professional career. Each of these scholarship holders receives 300€ jointly funded by public and private means (BMBF 2012a).

As the majority of German universities do not charge tuition fees from their students, merit-based stipends in Germany are rather aimed at identifying future top performers at a very early stage of their career (BMBF 2009, 2011). In addition to direct financial sponsorships, German merit-based scholarships also support their stipend awardees in non-material ways by granting them access to e.g. networks and elite positions (BMBF 2009). At the moment however, merit-based aid does not rate as high as need-based student aid in Germany. Only approximately 1% of all German students, i.e. 24,000 individuals, is sponsored by any of the above-mentioned organizations (BMBF 2009, 2012b; Middendorff, Isserstedt and Kandulla 2009). Federal funds provided for this purpose however have increased substantially from 80.5 million € in 2005 to 132.3 million € in 2009 (BMBF 2009) and the Federal Ministry of Education and Research aspires to increase this amount substantially in the coming years (BMBF 2011).

[5] These 12 organizations are: 'Cusanuswerk', 'Ernst Ludwig Ehrlich Studienwerk', 'Evangelisches Studienwerk e.V. Villigst', Friedrich-Ebert-Stiftung', 'Friedrich-Naumann-Stiftung', 'Hanns-Seidel-Stiftung', 'Hans-Böckler-Stiftung', 'Heinrich-Böll-Stiftung', 'Konrad-Adenauer-Stiftung', 'Rosa Luxemburg Stiftung', 'Stiftung der Deutschen Wirtschaft' as well as the 'Studienstiftung des Deutschen Volkes'. For more detailed information on each of these organizations see e.g. www.begabtenfoerderungswerke.de or BMBF (2009).

3 RELATED LITERATURE

In order to relate the investigation of who is awarded a scholarship to previous research activities, the following section gives an overview of the related literature. To the best of my knowledge, only very few researchers have previously explicitly investigated who succeeds in a stipend selection process. The only two investigations addressing an at least similar research question have been conducted by Freeman (2005) and Middendorff, Isserstedt and Kandulla (2009).

Freeman (2005) empirically investigated the determinants of stipend awardees for graduate students using panel data of the National Science Foundation (NSF) Graduate Fellowship Research Program. With the purpose of giving policy advice of how to raise the number of US natives choosing science and engineering careers, he empirically tested how several factors affected the probability of being awarded a stipend over a period of 22 years (1976 to 1998). Freeman (2005) finds scholastic achievements such as Grade Point Averages (GPA), Graduate Records Examination (GRE) scores and the quality of reference letters to be most important for the probability of receiving a stipend. Furthermore, controlling for demographic determinants revealed that – in line with the diversity criterion – women and minority group members c.p. have better chances of receiving the award than majority men. Investigating more than 100,000 award decisions over a period of more than 20 years, Freeman (2005) is able to explain a high proportion of variation in both award offers and panel ratings. Although originally aiming at a different purpose – giving policy recommendations of how to increase the number of native Science and Engineering (S & E) graduate students by increasing the number of stipends awarded – Freeman's (2005) paper gives a first insight into the determinants of stipend awarding. It was the first investigation of the criteria evaluators in a stipend selection process apply.

Middendorff, Isserstedt and Kandulla (2009) conducted the only study about stipend granting behavior in Germany, but went into a slightly different direction than Freeman (2005). With the help of a survey conducted in October 2008 among all 19,958 actual scholarship recipients of eleven German institutions who provide scholarships for outstanding students[6] they tried to find out how the social profile of an average student

[6] '*Begabtenförderungswerke'* in German. Note that of all 12 institutions presented in Chapter 2, the 'Ernst Ludwig Ehrlich Studienwerk' was not considered in the Middendorff, Isserstedt and Kandulla (2009) study.

scholarship recipient looks like. [7] In order to be able to classify and evaluate the findings of their investigation, Middendorff, Isserstedt and Kandulla (2009) consulted data from another survey based on a representative German student sample ('18. Sozialerhebung des Deutschen Studentenwerks', see Isserstedt et al. 2007 for details). Having analyzed the questionnaires of all stipend awardees who voluntarily took part (48% of all current scholarship holders), Middendorff, Isserstedt and Kandulla (2009) inter alia found evidence for some kind of social imbalance: The major part (51% resp. 21%) of scholarship holders comes from the upper (middle) class and children from less educated backgrounds are highly underrepresented (less than 10% have a working class background). The classification to these 'social groups of origin' is based on a combination of both the parents' educational background and current occupation. [8]

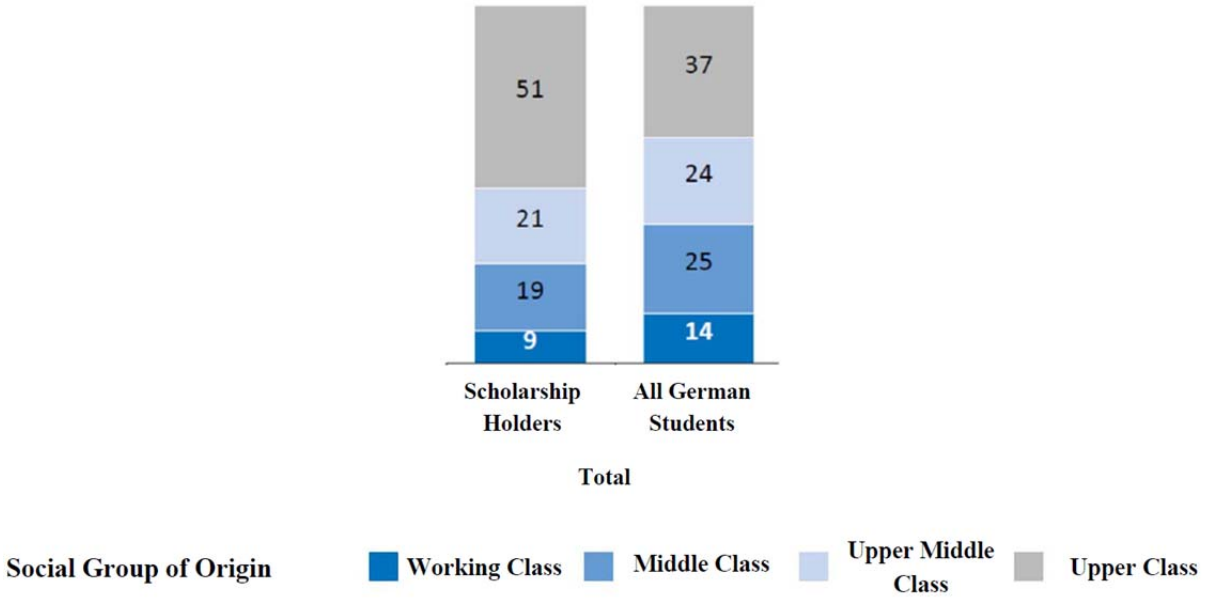

Figure 3-1 Social Origin of Stipend Awardees (left column) and All Students in Germany (right column) (Middendorff, Isserstedt and Kandulla 2009, 28).

Compared to all students in Germany, the fraction of those being classified as having an upper class background is significantly higher among sponsored students (51% vs. 37%), whereas students from the upper middle class (21% vs. 24%), the middle class (19% vs. 25%) and the working class (9% vs. 14%) are underrepresented in comparison to the

[7] Middendorff, Isserstedt and Kandulla (2009) divide their study into two sections: (undergraduate) student and doctoral stipends. Here, only the results concerning (undergraduate) student profiles are presented, as this is the population of interest in the present research. Students also represent the majority of all stipend awardees: approximately 85% of all 19,958 scholarship recipients in October 2008 were students (16,935 vs. 2,949 doctoral students, s.p. 14).

[8] For a detailed explanation, see Middendorff, Isserstedt and Kandulla (2009), p. 28-30 and Appendix, Picture A.1.

entire student body in Germany (see figure 3-1). In consequence, the authors postulate that a social injustice in scholarship provisions exists. These kinds of scholarships rather seem to be granted to those who have access to a good education anyways (even without the scholarship) and – following their assessment – fail to support those really in need of such a fellowship. Middendorff, Isserstedt and Kandulla (2009) also compare the fraction of students with a migrational background among sponsored and 'regular' students, but do not find any substantial differences between both stipend awardees and all students. They do find some further differences in the demographic profile: stipend awardees are for instance slightly younger and more frequently live outside of their parental home. Additionally, some academic, scholastic and professional characteristics also differ between the two groups (e.g. the fraction of students having completed an apprenticeship prior to their studies is significantly lower among stipend awardees and the percentage of those having already studied or lived abroad is significantly higher among sponsored students). Stipend awardees also work part-time less frequently and if they do, they rather do it for career instead of monetary reasons (as 'regular' students mostly declare). Nevertheless, the disparity in social backgrounds of stipend awardees and all students remains the most obvious result Middendorff, Isserstedt and Kandulla (2009) found in their study.

Even though this study is an important step towards understanding scholarship granting behavior, it lacks some important steps in order to be able to make inferences about how selection procedures in this context work. One major shortcoming of this investigation is that it addresses only those who have been granted the scholarship. Although getting in-depth information about the background of those who are financially supported by one of these institutions, we have no information about those who were rejected. We are therefore not able to find out whether e.g. the apparent uneven social distribution among scholarship recipients is due to self-selection processes (i.e. because only those having an upper class background applied for the scholarship) or whether evaluators discriminate against applicants with a working class background. This deficit can only be eliminated by investigating the entire selection process as opposed to only asking the successful applicants some questions about their social background.

Apart from the above mentioned studies, no (empirical) research has been published dealing with the determinants of stipend awards.[9] Only in laboratory experiments, a fictitious setting of the selection of scholarship applicants has been used to demonstrate and explain physical attractiveness and similarity-attraction phenomena (Agthe, Spörrle and Maner 2010).[10]

Due to the lack of further literature on (merit-based) scholarship awarding decisions, literature dealing with related fields needs to be used as a reference for hypotheses development. First, the scope of related literature will be extended to include empirical elite research in general and upward mobility in particular in Chapter 3.1. Understanding a) who is considered to be an elite member and b) who is actually able to move up to elite status in Germany is crucial for getting an impression of how elites are built in Germany. Only then it is possible to anticipate which characteristics and attributes might influence the evaluator's decision in stipend awarding decisions.

Subsequently, the focus of the literature review turns to success factors in other types of selection processes. First, some papers investigating determinants of success in an educational selection process need to be discussed (Chapter 3.2). Finally, most important findings for the type of selection process that has been studied most frequently so far – the personnel selection process – will be presented in Chapter 3.3, arranged according to the respective influence factor(s) that have been studied in these papers.

[9] All other studies on scholarship programs that have been published so far rather deal with the optimal financial design of scholarship programs in order to maximize either the number of students going abroad (Lien 2007) or their post-program ability (Lien Liu 2010; Lien Wang 2010) and not with the question of who is actually selected. Additionally, some empirical research on the impact of merit-based financial assistance on college enrollment has been conducted (Cornwell, Mustard and Sridhar 2006; Monks 2009; Stanley and French 2009). Furthermore, self-selection into, i.e. application for, a (means-tested) scholarship program in the United States has been empirically tested by Singell (2002) and Figlio, Hart and Metzger (2010). Mead (1965) gives a descriptive overview of a union scholarship program and Opheim (2006) examines policy-induced changes in the number and composition of (means-tested) student support in Norway. But as outlined in Chapter 2, merit-based and means-tested scholarships serve completely different purposes and therefore, totally different admission criteria will be applied by the evaluators for each type of scholarship program.

[10] A more detailed presentation of the results of this study is provided in Chapter 3.3.2.

3.1 WHO IS ABLE TO BECOME AN ELITE MEMBER IN GERMANY?

First and foremost, it is not my intention to give a comprehensive (historical) literature review on German empirical elite research.[11] In this section, only a few selected empirical investigations dealing with the question of who belongs to the German elite will be presented, particularly focusing on the question which of an individual's attributes are decisive for elite recruitment.

In elite research, there is a lack of unity in the understanding of elite formation. According to Kaina (2006, 2009), there is still no general consensus about who belongs to the elite(s) of a particular society and why. Empirical research on elite formation mainly relies on characteristics of those who are currently understood as being elite members.[12] With the help of detailed surveys and interviews, elite researchers inter alia try to find out which attributes are of importance in elite recruitment. The most important empirical studies in Germany trying to find answers to this question are three surveys known as 'Mannheimer Elitestudien' which were conducted in West Germany in 1968, 1972 and 1981 respectively as well as a follow-up study called 'Potsdamer Elitestudie' in 1995 in which executives from the reunified Federal Republic of Germany were interviewed. Another important analysis has been conducted by Hartmann (2002) who tracked and analyzed both the origin and the career progression of four cohorts of PhD graduates in Engineering, Law and Economics. The results of the two most recent elite studies (1981 and 1995) will be presented briefly and compared to what Hartmann (2002) found in his investigation in the following sections.

[11] Readers interested in a general overview: see e.g. Hoffmann-Lange (1992; 2001), Hartmann (2002) and Kaina (2009).

[12] Who is to be interviewed in these surveys, i.e. who is considered to be an elite member, can be elaborated following three different approaches (reputational approach, decisional approach, positional approach). For a detailed explanation of these three elite member identification approaches, see e.g. Putnam (1976), Felber (1986), Hoffmann-Lange (1987, 2007) or Kaina (2009).

Procedure of the Elite Studies in 1981 and 1995

All of the four main German elite studies strived for a census of all current occupants of leadership positions in Germany (Machatzke 1997). In order to be able to conduct comparisons between all elite studies, maximal consistency in terms of both content and selection of respondents was intended. Respondents in all four studies were selected following the positional approach which is the most frequently used approach in the social sciences: This approach is the simplest (and therefore fastest) of all approaches and is based on the assumption that power in industrialized societies is linked to formal positions of leadership (Wildenmann et al. 1982; Hoffmann-Lange 2007) rather than to individuals. According to this approach, several positional elites are determined following a multi-stage procedure: First, all relevant sectors of society are identified. Second, within each sector, influential organizations such as political parties or corporations are identified. Third, executive positions are determined and current occupants of these positions are contacted. Being the easiest and most reliable way of identifying the target group, the positional approach is the most frequently used method in the social sciences (Kaina 2009).

Sectors of society that have been included in the Mannheim Elite Study in 1981[13] (Wildenmann et al. 1982) were

(1) Politics
(2) Administration
(3) Business (Commercial Enterprises)
(4) Trade Associations
(5) Unions
(6) Mass Media
(7) Science
(8) Military
(9) Culture
(10) Other

[13] Due to the longitudinal design, basically the same sectors have been included in the remaining three elite studies.

In total, 3,580 positions had been identified in 1981 (Wildenmann et al. 1982). In 1995, the number of executive positions investigated totaled 4,569 (von Rosenblatt et al. 1995). Due to vacancies, job rotations during the observation period or accumulations of multiple positions, the number of actually contacted persons occupying these positions was reduced to 3,165 in 1981 and 3,941 in 1995. All of the identified subjects were informed in writing about the study and its purpose and were asked to participate in individual interviews.[14] Table 3-1 illustrates the number of individuals contacted in each year, the response rates as well as the final number of participants in each sector.

	1981			1995		
Sector	Contacted Subjects	Response Rate	Subjects Interviewed	Contacted Subjects	Response Rate	Subjects Interviewed
Politics	452	60,6%	274	898	55,6%	499
Administration	471	62,8%	296	646	73,4%	474
Business	688	41,4%	285	651	38,2%	249
Trade associations	296	58,8%	174	310	55,8%	173
Unions	155	56,1%	87	164	59,1%	97
Mass Media	354	62,7%	222	454	61,9%	281
Science	179	72,6%	130	202	81,2%	164
Military	172	25,0%	43	157	86,0%	135
Culture	180	57,8%	104	178	56,7%	101
Other	218	59,2%	129	281	69,8%	168
Total	3,165	55,1%	1,744	3,941	59,4%	2,341

Table 3-1 Number of Contacted Subjects and Response Rate in the Elite Studies of 1981 and 1995
Source: Wildenmann et al. (1982), 16-17 and von Rosenbladt et al. (1995), 33.

The data analysis mainly remained descriptive: Proportions of several groups in both the elite and the entire population at this particular point of time were compared and (dis)proportionality indices in several sectors were determined. In order to find out whether or not access to elite positions is determined by social background or education or both, also logistic regression models were used to isolate these effects.

[14] Detailed information about all questions asked during the interviews can be obtained from the appendices of Wildenmann et al. (1982), Hoffmann-Lange (1992) and Bürklin et al. (1997). The original questionnaires of both studies can also be requested from the 'GESIS - Leibniz Institute for the Social Sciences/Data Archive for the Social Sciences (DAS) Germany' (http://www.gesis.org/en/services/research/data-catalogue/) under the heading 'ZA1139' for the 1981 Mannheim Elite Study and 'ZA2881' for the 1995 Potsdam Elite Study.

Procedure of the Hartmann study

Hartmann (2002) chose a different approach. Believing that the doctorate is the highest (general) educational qualification, Hartmann (2002) assumes that one will find the highest social selectivity among PhD graduates. He also denotes that a PhD in the fields of Engineering, Law and Economics is disproportionately common among elite members, and therefore examines only career paths of four cohorts of PhD graduates in these fields. Hartmann (2002) chose to track the careers of 6,500 individuals who graduated in either 1955, 1965, 1975 or 1985. In order to be able to analyze both origin and career development of these individuals, he extracted important variables of the graduates' CV.[15] Additionally, Hartmann consulted relevant compendia (e.g. the 'Hoppenstedt-Handbuch' for business careers) approximately ten years after each graduation in order to add occupational information to the data set. Apart from only including PhD graduates, the main difference to the traditional elite studies is that Hartmann does not interview his subjects, but tries to consult more objective data from official sources.

Hartmann (2002) also mainly illustrates his findings descriptively using separate proportions for e.g. different social classes, but additionally applies more sophisticated multivariate procedures such as Event History Analysis to find answers to how access to and success in elite positions is determined.

Results

The fact that access to elites mainly depends on the social background of an individual has been widely acknowledged by all researchers using the data of the above mentioned four studies (Hoffmann-Lange 1992; Schnapp 1997; Hartmann 2002, 2004, 2007): Children from upper (middle) classes have relatively better opportunities to reach an elite position than individuals possessing 'only' a middle class background. But this social imbalance does not hold true for all elite sectors: Whereas access to business, public administration and military elites is mainly granted to upper class members, elite positions in law, science, union and political fields are filled with individuals from all social classes. Being aware of this (partial) social imbalance in elite recruitment and making use of the quasi-longitudinal data basis, also changes in elite composition over time can be examined: Has the relative influence of social origin decreased over time? As functional elite theory predicts,[16] have meritocratic, i.e. performance-related, selection

[15] Age, gender, degree-granting university, field of studies, beginning & length of study, number and location of universities attended, full/part-time employment before, during & after the studies as well as the social origin (3 categories) count among the variables Hartmann (2002) was able to extract from most of the CVs attached to the PhD thesis. For a more detailed description of his research design, see Hartmann (2002), 31-43 & Hartmann (2001), Chapter 2.

[16] For a detailed discussion of functional elite theory, see Chapter 4.1.2 of this thesis.

criteria gained in importance over the last decades? Schnapp (1997), Rebenstorf (1997) and Hartmann (2002) draw quite controversial conclusions when addressing this issue: Whereas Hartmann (2002) even denotes tendencies for more severe social imbalances in the recent past in his study, both Schnapp (1997) and Rebenstorf (1997) provide support for a (slightly) facilitated, i.e. broader, access to elite positions compared to the results from the earlier 'Mannheimer Elitestudien'. These differences might occur due to the different methodological strategies applied. Whereas Hartmann (2002) only considers official information about the careers of PhD graduates, the traditional elite studies follow the positional approach and interview persons in elite positions regardless of their education. Both strategies have inherent advantages and disadvantages, but as a result of the differing sample they can lead to different results. Secondly, the classification of the social background also varies among the studies. The elite study uses Goldthorpe's (1982) concept of the service class, while Hartmann (2002) originally built eleven categories and combined these to three social classes which are not necessarily congruent with the service class classification.[17]

The question remains whether the social background has a direct or only an indirect effect (or even both) on elite access opportunities. Schnapp (1997) illustrates all possible relations between the three variables 'Social Background', 'Education' and 'Elite Status' in four different models (see figure 3-2).

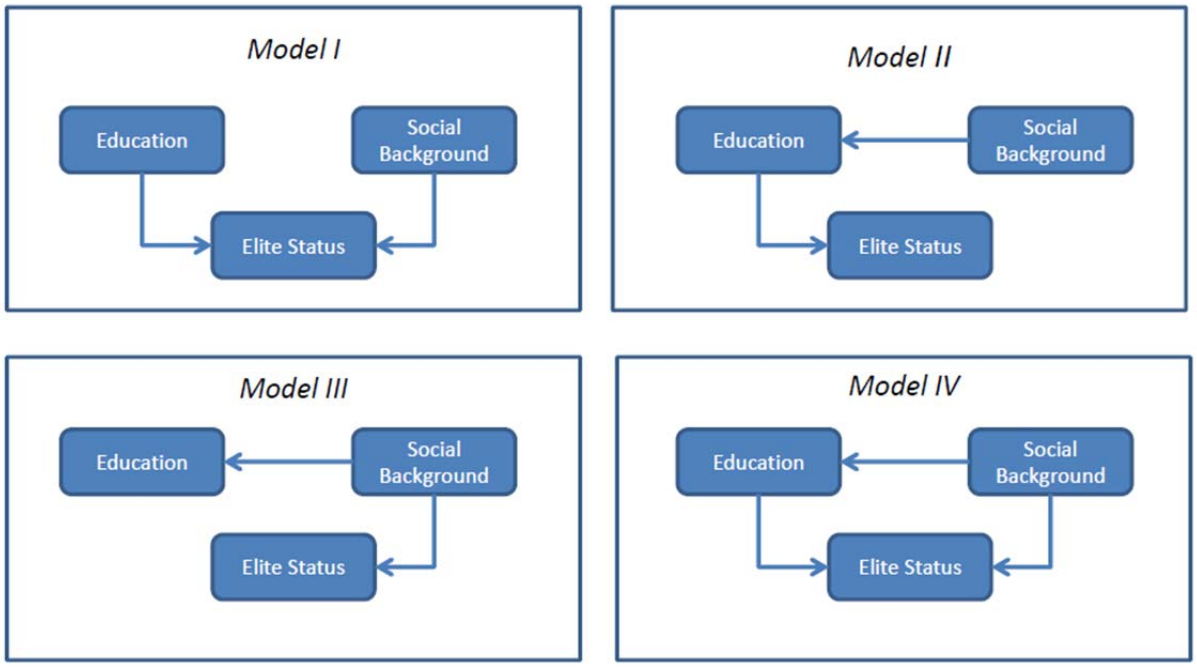

Figure 3-2 Models Reflecting the Relation between Social Background, Education and Elite Status (according to Schnapp 1997, 84).

[17] For a more detailed explanation of how both authors classified the social background, see Hartmann (2002), 33-34 and Schnapp (1997), 72-73.

It is feasible that both meritocratic (education) and non-meritocratic (social background) attributes can have an influence on whether or not somebody belongs to an elite. Whenever each variable has a direct separate influence on elite status, but both effects are independent from each other, Model I would be supported. Modell II by contrast postulates only a direct effect of somebody's background on his or her course of education. Having achieved a better education then in turn improves the chances of becoming an elite member, but no direct effects of somebody's social background can be found. This would mean that the social background only indirectly affects elite access as it facilitates access to education which is a prerequisite for elite status. Following Model III, one would expect a direct effect of social background on both education possibilities and elite admission, whereas Model IV claims that social background has both a direct and an indirect effect (via education) on elite membership.

As e.g. Schnapp (1997) concludes, higher educational achievement, i.e. a university degree, indeed leads to improved recruitment chances. But as the probability of achieving a university degree in Germany is significantly higher for upper class members,[18] this might be the actual driver of unequal elite access opportunities (as Models II and IV predict). But both Hartmann (2002) and Schnapp (1997) further observe a direct influence of a person's social background although controlling for given unequal chances to obtain a university degree. This leads to a 'double privilege' for upper class members which is reflected in Model IV and may be explained by the use of Bourdieu's (1983) different kinds of capital. Due to cultural and economic capital, descendants from upper class families face better chances to obtain tertiary education and subsequently a university degree. Apart from that, they also possess higher social capital which basically means that they have access to social networks allowing them to inofficially gather information that might stimulate their career (Schnapp 1997). This lack of social capital can only be partly compensated for by acquiring a good tertiary education and individuals coming from lower social classes can thereby only partially improve their recruitment opportunities (Schnapp 1997).

Apart from the social background of a person, other socio-demographic characteristics were also salient in the elite populations from 1981 and 1995 and differ significantly from the entire population (see table 3-2).

[18] See e.g. Bertelsmann Stiftung (2012).

	1981		1995	
Socio-demographic Attributes	**Elite Members**	**German Population**	**Elite Members**	**German Population**
Age	53	38	52	40
Proportion of Females	3%	52%	13%	51%
University Degree	68%	4,7%	78%	7,6%
Thereof:				
PhD	49%	n.a.[19]	47%	n.a.
Habilitation	8%	n.a.	8%	n.a.
Total Subjects Considered	1,744	61,658,000	3,241	81,817,000

Table 3-2 Selected Socio-Demographic Characteristics of Elite Members in 1981 and 1995
Sources: Values for Elites 1981: own calculations based on the data set of the Mannheim Elite Study (ZA1139); Values for Elites 1995: Kaina 2009, 402. German population indices calculated from Destatis 2012a, 15-17; Destatis 2011a, 12.

Overall, elite members are relatively old in all four studies: In 1995, they were on average 52 years old (Kaina 2009), but the average age varied significantly in different elite sectors. Whereas elite members in politics and mass media have always been the youngest, elite positions in business, administration and trade associations are constantly filled with the oldest of all elite members. This is directly related to the fact that career advancement is mainly driven by seniority: Especially in West Germany (in 1995), most of the elite members had been working for this specific organization for more than a decade before advancing to the current elite position with an average age of 49. Consulting the findings of Opitz (2005) on the career progression of top managers in the US, France and Germany, it is not likely that this finding represents a peculiarity of the German elite. Transferring Opitzes (2005) results for high potentials on elites, German elite members are not expected to be significantly older than their international equivalents. Women are traditionally underrepresented in all elite positions (Kaina 2009), although the overall percentage of female executives has constantly risen from 2% in 1968 to 13% in 1995. Especially, in business (1%), trade associations (2%), science (3%) and military elites (1%), women were extremely underrepresented in the 1995 study whereas female elite members in politics account for more than a third of all executives (36%). To conclude, elite positions are still dominated by relatively old, male and highly educated individuals stemming from the upper classes. Nevertheless, the survey results from 1995 indicate a slight tendency towards broader access opportunities for lower class members as long as they take the opportunity to obtain the same tertiary education as their upper class 'competitors'.

[19] Prior to 1998, the PhD was not separately recorded by the Federal Bureau of Statistics ('Statistisches Bundesamt') and was simply classified as 'university degree'. Only data on all PhD graduations in a given year is available for these years (Destatis 2011b, 10).

3.2 Who Succeeds in Other Educational Selection Processes?

In the absence of empirical findings (or better: publications) on scholarship awarding decisions, it might be helpful to look at other selection processes in educational settings. Although knowing that the purpose of these selection processes – and therefore also the 'ideal' candidate - can be quite different from the current context, some success factors identified in other educational selection processes may also prevail in stipend granting behavior and may assist in hypotheses development for the current study.

When it comes to selection processes in education, college admission decisions are one of the most frequently analyzed subjects: Who is admitted to a certain college and who is rejected? Which are the selection criteria colleges base their decision on? How do colleges screen prospective students? How do college applicants signal their aptitude for studying at the college/university they applied for? The most important findings of papers trying to find answers to these questions will be presented subsequently.[20]

Facing an increasing supply-demand imbalance, colleges and universities all over the world need to establish fair, consistent, reliable and valid selection processes (Turner and Nicholson 2011) in order to minimize misjudgments. In Germany, university admission decisions are traditionally mainly based on the high school grade average[21] while the use of other criteria such as standardized aptitude tests is rather exceptional (Formazin et al. 2011). In the United States and the United Kingdom however, college admission is quite selective[22] and the aptitude of applicants needs to be assessed in advance. As colleges and universities lack a uniform definition of aptitude and as academic achievements always have to be assessed in the educational context of an applicant (Stringer 2008), selection or rejection decisions are based on several meritocratic and non-meritocratic criteria. How these criteria look like has for instance been empirically tested in several studies in the UK, the Netherlands and in the US. Numerous studies on college admission have been conducted in medical or psychology school contexts which seem to be the most selective fields of study (in Europe) and consequently much attention is placed on the selection of medical students (Carr 2009). Aspiring to select future doctors in a more holistic way than solely looking at academic qualifications (Nicholson and Turner 2011), British medical schools select with the help of diverse criteria provided by students in their applications, especially in times of

[20] As being admitted to a higher education institution is a prerequisite for having the chance of being awarded a study-related scholarship, success factors of college applications may also be of importance to stipend awarding decisions.

[21] For students who do not meet the required numerus clausus at once, waiting periods improve the chances of being admitted to university (www.wartesemester.de).

[22] At least for colleges or universities of high quality (elite institutions).

decreasing academic variability. Formal aptitude testing has only recently been introduced in the UK as a new means of selecting prospective students and is still subject to controversial public discussion (McDonald et al. 2000). High school graduates wishing to attend medical school hand in their application form to a central body, the Universities and Colleges Admission Service (UCAS). Completed application forms include personal information about the candidates, details about their academic achievements, a letter of motivation for their wish to study medicine as well as at least one letter of reference. The statements provided in these forms are usually verified by several attached official documents and certificates (Nicholson and Turner 2011). Selectors at medical schools screen these forms and identify which candidates they want to reject at this early stage and which of the applicants they would like to invite to a personal interview.

With the use of both a focus group design with 17 selectors of a London based medical school as well as a content analysis of selector marking forms, Nicholson and Turner (2011) identified criteria frequently used for both selection and rejection of a candidate. Most frequently, the provision and reflection of medically-related work experience led to the invitation of a candidate. Although there was only little consensus about the appropriate or minimum amount of work experience a candidate needs to show, all focus group participants agreed that a lack of any work experience in the medical field leads to the immediate rejection of a candidate. Furthermore, showing some commitment of any sort – extracurricular activities, sports or even a part-time job on weekends – revealed to be advantageous for an application to be considered. The same applies for personal statements that have to be judged quite subjectively: As long as the candidate is able to show any interest outside the academic field and conveys his or her ability to work in a team, personal statements rather lead to the selection of candidates. Negative teacher references are seen as alert signals and often lead to a rejection of a candidate. (Overly) positive letters of recommendation in turn do not necessarily lead to an invitation of the candidate, as selectors do not perceive these letters as being explicit enough as they have to 'read between the lines'. This corresponds to the limited predictive power of teacher references that Ferguson et al. (2003) found when examining the relationship between personality, references and personal statements with performance in medical school. "In summary the 'ideal' candidate was selected for having undertaken and reflected upon an appropriate amount and type of medically-related work experience, having a supportive teacher reference, possessing positive attributes detailed in the candidate's personal statement, and demonstrating commitment to the study of medicine" (Nicholson and Turner 2011, 305).

The advantageousness of such costly and time-consuming selection procedures in comparison to selection by lottery has been shown in the Netherlands by Urlings-Strop et al. (2009). In a controlled experiment they identified that assessing applicants' cognitive and non-cognitive aptitude a priori leads to the selection of students whose dropout rate is significantly lower (2.6 times) than that of students who have been admitted by lottery. In terms of academic performance however, no significant differences have been found between these two groups.

For the United States, different college admission criteria might apply. In contrast to most European countries, formal aptitude testing (SAT) has had a long tradition in the US. Nevertheless, at least for some high quality colleges and universities, some factors other than SAT scores play an important role in admission. For psychology graduate students, Rodolfa et al. (1999) have identified several exclusion and inclusion criteria in the predoctoral internship selection process which represents the last step for graduate students before earning their doctoral degree. Looking at graduate students' selection success might be particularly beneficial for the study of scholarship success factors, as applicants for stipends are usually also advanced students. The factors that selectors indicated in the 1999 survey to be most important for selection are in descending order (1) applicant fit, (2) supervised clinical work experience, (3) completion of related doctoral coursework, (4) the interview, (5) the status/reputation of the attended doctoral program, (6) completion of written and oral examinations, (7) the professional demeanor displayed by the applicant and (8) the letter of recommendation provided by the applicant. Distinct exclusion criteria are the lack of accreditation of the doctoral program and the non-completion of exams. Ginkel, Davis and Michael (2010) replicated the survey after the introduction of a standardized application procedure and mainly found personal characteristics – as opposed to meritocratic attributes – to have increased in importance. They interpret the change as a way for selectors to differentiate between applicants of equal academic achievements.

Apart from peculiarities of medical and psychological curricula, other American studies have tried to identify more general US college admission criteria for diverse disciplines. One of the most comprehensive investigations trying to identify how US institutions select undergraduate students has been conducted by the nonprofit College Board Association (Rigol 2003). With the help of interviews, site visits and examining both official and internal college materials, Rigol (2003) found that there does not seem to be a best practice of how college admission decisions are made in the United States. Various heterogeneous approaches to student selection exist and each college or university has

elaborated individual policies and practices. Depending on the institutional mission pursued and therefore the desired outcome for its students, colleges adopt different selection strategies. An approach which is commonly adopted by public colleges and universities is to admit all students meeting certain predetermined requirements such as grades or test scores above a certain threshold. These institutions are called entitlement or open access institutions. Other institutions rather want to maximize the success of admitted students and admission mainly depends on how these colleges or universities define success. These types of institutions look for a student body that optimally reflects the institution's vision. This does not necessarily mean that each individual student needs to meet certain predefined criteria, but the best balance of student skills, talents, backgrounds and interests is sought after. Naturally, all selection approaches are subject to constant change according to supply and demand in a given selection year. In her report, Rigol (2003) offers an in-depth analysis of the different approaches higher education institutions in the US adopt and outlines seven different models that are frequently used to select undergraduate students. Although differing very much in their selection practices, most colleges and universities share the use of the following categories of factors considered (Rigol 2003, 19):

a) **Academic Achievement, Quality and Potential**
   Direct Measures
   Caliber of High School
   Evaluative Measures
b) **Nonacademic Characteristics and Attributes**
   Geographic
   Personal Background
   Extracurricular Activities, Service and Leadership
   Personal Attributes
   Extenuating Circumstances
   Other[23]

As this list shows, admission criteria in the US are influenced both by meritocratic and non-meritocratic characteristics.

[23] A summary of application components is provided by Rigol (2003), 61-67 and a more detailed enumeration of all factors that may be used in selection can be found in Rigol (2003), 75-77.

One of the largest and most frequently cited scientific US studies on college admission has been conducted by Manski and Wise (1983). In the course of empirically investigating individual economic higher education decisions and behavior of US citizens, Manski and Wise (1983) also examined the admission behavior of universities and colleges of different quality. By simultaneously looking at individual application and institutional admission decisions, Manski and Wise (1983) found that on average, the admission to a four-year college of a certain quality is mainly influenced by the individual decision to apply for it. This means, college admission is rather the result of applicant self-selection than a consequence of admission officers' decisions. Using data from the National Longitudinal Study of the high school class of 1972, Manski and Wise (1983) conducted a multivariate analysis comparing the effects of four groups of variables measuring academic potential as well as nonacademic attributes on both application and admission to colleges of different qualities.[24] Being able to use such a broad dataset encompassing all high school graduates, they do not only have information about actual college applicants, but also about those who decided not to go to college.

The four groups of variables Manski and Wise (1983) tested included

(1) academic and nonacademic high school performance,
(2) an applicant's socio-economic background,
(3) local labor market conditions representing alternatives to studying and
(4) indicators of sex and high school environment (urban vs. rural).

The results show that the "most important determinant of both college choice and admission is scholastic preparation, as reflected in the SAT score and high school class rank measures" (Manski and Wise 1983, 84). For college admission, college quality represents the key factor: Colleges already hosting better performing students seem to be more selective than colleges requiring only lower academic aptitude and admitting freshmen with lower SAT scores. Therefore, SAT scores are an important determinant of selection outcomes of high quality colleges. Leadership positions as well as athletic achievements in high school only have a minor influence on college admission decisions, but do affect individual application decisions significantly. Concerning the socioeconomic background, parents' education and income positively affects college application and college quality. As the returns to college education are generally higher for black high school graduates, they are c.p. more likely to apply to a college than Whites. For admission decisions however, the race of an applicant is irrelevant.

[24] School quality is measured as the average combined SAT score of freshmen entering the school to which a student applies (Manski and Wise 1983, 69).

Manski and Wise (1983) were able to show that measures of academic aptitude such as SAT scores are much more important for individual application decisions than for institutional admission results. This reflects a substantial influence of self-selection in higher education decisions: “Although people with low academic ability and poor past performance are very unlikely to apply to any four-year college, such people, if they were to apply to a college of average quality, would have a high probability of admission. The probability of admission is much higher than the probability of application” (Manski and Wise 1983, 89).

In summary, the way in which signaling and screening in college admission decisions work does not only vary between countries or continents, but also within a country, very different selection criteria are applied. Depending on the status and the quality of an institution, totally different admission criteria emerge. But in order to ensure a selection process as transparent and fair as possible, most institutions mainly rely on ‘objective’ criteria such as test scores, high school performance, work experience and references.

Applying these results to the present situation on the one hand assists in understanding how selection processes in educational settings work in general. But on the other hand, one needs to bear in mind that the rationale behind student selection and/or admission is quite different from stipend awarding decisions: Whereas colleges and universities try to maximize average student success – however this might be measured –, scholarship granting organizations aim at supporting future elite members in both material and non-material ways (BMBF 2009). Due to this goal discrepancy, success factors in stipend awarding decisions might vary widely from college admission success factors.

3.3 What Can We Learn from Personnel Selection Processes?

In the absence of further empirical investigations of related elite or educational selection criteria, it may be worthwhile looking at success factors of other applicant selection processes being characterized by a substantive supply-demand imbalance. The most extensively investigated applicant selection process in economics, psychology and social sciences is the personnel selection process. As choosing the wrong candidate can become very costly, a large interest in the drivers and determinants of personnel selection exists. Research questions commonly addressed are: Who among a job applicant pool gets the opportunity to present himself in an interview? Who in the end gets the job? Which are credible signals job candidates provide in their application? How do employers screen the ability of applicants? Are there any factors prevalent in the selection process that are of rather subjective nature and bias the evaluators' decision? Numerous researchers focusing on the strand of research dealing with "explaining and predicting an individual's success in job search" (Chia 2005, 75) have addressed these questions empirically so far. Studies trying to reveal the relative importance of various factors on somebody's selection success can be classified into three major categories according to the method used in data collection. Researchers either

(1) conducted a survey among (potential) employers asking them to indicate the relative importance of several criteria in selecting a new employee,
(2) reviewed actual success rates of real applicants, or
(3) performed experiments by manipulating two or more independent variables.

Additionally, meta-analyses try to summarize the respective effects of a certain influencing factor found in numerous previous studies in diverse fields and occupations. All of the above mentioned methods do have several respective limitations that need to be mentioned. Data gathered from questionnaires sent to employers will always be subject to manipulation by those filling in the survey (stated preferences): This may result in subjective and sometimes even false evaluations of selection criteria that do not reflect reality as e.g. Behrenz (2001) and Cole et al. (2007) were able to show. Actual success criteria from real selection periods (revealed preferences) are hard to get as most employers usually do not publish the characteristics of either applicants or newly recruited employees. Experiments in turn only represent a very simplified model of reality in which the applicants only differ in two or three key characteristics/attributes. The most important findings from all these kinds of studies will be presented in the following sections.

Due to the abundance of empirical papers addressing personnel selection success factors, the results will be arranged in the order of the influence factors examined in the respective studies. Whenever the importance of a certain factor of influence varies according to the screening activity employed by the recruiter (e.g. pre-selection based on paper credentials vs. final selection after an interview), the respective effects will be discussed separately. Beginning with rather objective criteria that can be classified as signals (Chapter 3.3.1) according to Spence (1973)[25], the review subsequently turns to the respective influence of factors that Posthuma, Morgeson and Campion (2002) label individual difference factors (Chapter 3.3.2) and situational factors (Chapter 3.3.3).

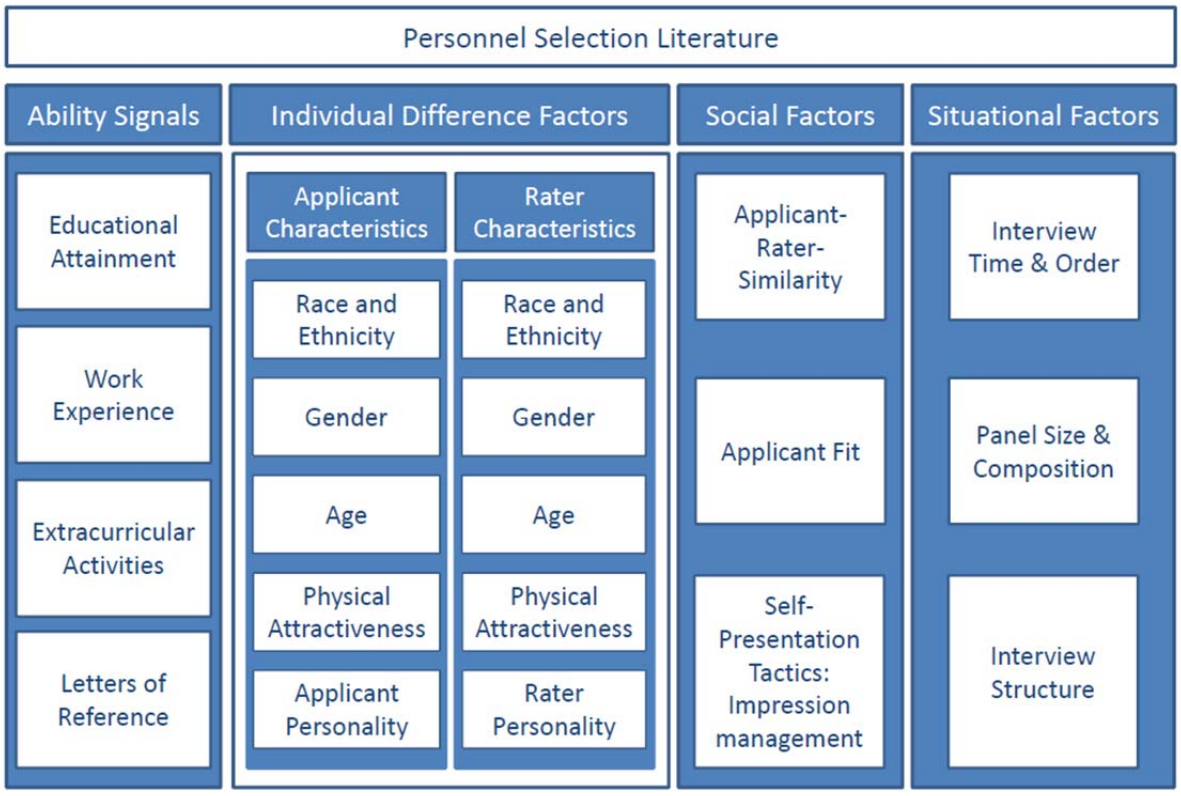

Figure 3-3 Structure of Chapter 3.3 (Personnel Selection Literature)
Source: Own Illustration

[25] For a more detailed explanation of Spences (1973) model, see Chapter 4.2.2.

3.3.1 The Influence of Ability Signals

Efficient and fair personnel selection should be based on objective, i.e. merit-based or job-relevant qualifications. In Spence's (1973) terminology, everything that can be classified as an objective measure of and credential for ability or future productivity is understood as a signal. In contrast to indices, signals are alterable by the individual applicant and therefore follow a function of the applicant's investment or effort. Recruiters commonly extract these kinds of ability signals from biographical data – also referred to as biodata - applicants provide in their application. Biodata is defined by Brown and Campion (1994, 897) as "work experience, education, activities, and other life history information contained in resumes and applications". With the help of this information, recruiters make inferences about a person's ability when screening (Brown and Campion 1994) and follow the rationale that "nothing predicts future performance as well as past performance" (Harold, McFarland and Weekley 2006, 337). In the aggregate, biodata inventories count among the most effective predictors of job performance (Brown and Campion 1994; Becton et al. 2009) and therefore represent a relatively valid selection device (Harold, McFarland and Weekley 2006).

Cole et al. (2007) were able to show that three résumé categories – academic qualifications, work experience and extracurricular activities (ECAs) – mainly predict recruiters' perceptions of an applicant's employability. Together with letters of recommendation (LORs) – which applicants can also indirectly improve via their effort and behavior – these signals should ideally have the greatest impact on employers' hiring recommendations as they represent the most objective measures of applicant quality.

3.3.1.1 Educational Attainment

Education is assumed to have a substantial impact on employers' evaluations of a job candidate. The influence of a person's education on employer ratings can be empirically tested using various independent variables: Either the influence of somebody's level of education – measured in terms of years of formal schooling or educational degrees obtained – or the respective absolute or relative performance in educational settings (e.g. high school grade average and class rank, university GPA and quartile rank in college graduating class, Dean's list and other awards etc.) can be tested.

Several studies and meta-analyses have shown that the validity of educational attainment and/or grades for adult achievement remains comparably low (Schick and Kunnecke 1981; Reilly and Chao 1982; Cohen 1984; Roth et al. 1996). Although academics agree on the low predictive value of grades for future (job) performance, many employers believe grades to be useful predictors of an individual's value to the firm (Reilly and Chao 1982; Roth et al. 1996). The main reason for the frequent use of grades and other educational achievement variables might be the immediate availability. Applicants usually hand in all the necessary information with their résumés and employers are able to use these criteria for (pre-) screening without additional costs. Although a direct ability measure such as cognitive ability testing would yield more valid predictions of future success (Berry, Gruys and Sackett 2006; Koedel and Tyhurst 2012), employers do use signals of educational performance very frequently. Believing that educational attainment is a proxy for cognitive ability (Berry, Gruys and Sackett 2006) and reflects desirable attributes such as intelligence and motivation (Roth and Bobko 2000), they refrain from using costly cognitive ability testing and use the information readily available in the résumé.

The most important findings resulting from surveys among actual recruiters (Lewis, Shimerda and Graham 1983; Behrenz 2001), reviews of actual success rates of real applicants with varying educational performance levels (Roth and Bobko 2000; Behrenz 2001; McKinney et al. 2003; Chia 2005; Berry, Gruys and Sackett 2006; Athey et al. 2007) and experiments having manipulated the level of academic qualification (Hakel, Dobmeyer and Dunnette 1970; Dipboye, Fromkin and Wiback 1975; Dipboye, Arvey and Terpstra 1977; Zikmund, Hitt and Pickens 1978; Knouse 1994a; Cole et al. 2007, Koedel and Tyhurst 2012) will be presented here. Additionally, selected results from meta-analyses (e.g. Cohen 1984; Olian, Schwab and Haberfeld 1988; Roth et al. 1996) will be discussed, too.

One major shortcoming of inferences made from experiments needs to be mentioned in this context: Very rarely, academic performance is the attribute of interest in experiments on personnel selection decisions. Educational attainment is rather used as a control variable reflecting applicant qualification while trying to reveal for instance discriminatory behavior of recruiters. In this function, academic achievement and work-experience are often manipulated as a combined measure of applicant qualification (e.g. Dipboye, Arvey and Terpstra 1977; Watkins and Johnston 2000) and single effects of academic achievements are often difficult to disentangle from work experience effects

(Cole et al. 2007). This occasionally exacerbates the interpretation of meta-analyses summarizing the influence of what is considered as 'qualification'.

Most studies provide support for a substantial positive influence of educational attainment, at least in personnel pre-selection or prescreening decisions. Only McKinney et al. (2003) found that 42% of the investigated screening decisions made by college recruiters were not at all influenced by the applicant's GPA. In comparison to the influence of several personal characteristics, objective qualification measures such as work experience and academic achievement are able to explain far more variation in selection decisions: Olian, Schwab and Haberfeld (1988) for example report that experience and education account for 35% of the variance in hiring recommendations, whereas individual difference factors such as gender are only able to explain single-digit percentages. Renwick and Tosi (1978) found that the undergraduate major and the graduate degree play a more influential role in selection decisions than personal characteristics. When comparing the relative influences of scholastic standing, work experience and extracurricular activities (ECAs), Hakel, Dobmeyer and Dunnette (1970) found evidence for an overwhelming reliance among interviewers on information about scholastic standing of recent accounting graduates: scholastic standing accounted for approximately 47% of the total variance in suitability ratings whereas the influence of both business experience and interests of the applicant remained negligible.

Both the duration of education and measures reflecting the performance in education positively influence the probability of being selected by recruiters. But the influence of grades varies widely according to the selection stage examined. Whereas measures of educational attainment are revealed to be a primary factor in initial screening decisions commonly leading to an invitation to an interview (Dipboye, Fromkin and Wiback 1975; Behrenz 2001; Chia 2005; Cole et al. 2007), their influence decreases or even disappears in later stages such as subsequent interviews or final job offers (Behrenz 2001; Chia 2005). In interviews, professional demeanor and what is commonly referred to as 'soft skills' become much more important than grades. As Harvey et al. (1997, Chapter 4) formulate, "having a degree … [is] … a necessary, but by no means a sufficient, criterion for getting a job […]. For many senior managers getting the job depends on such things, such as motivation and 'managerial potential' ". Nevertheless, Singer and Bruhns (1991) for example were able to show a substantial influence of higher levels of academic qualifications (MBA vs. Bachelor vs. high school certificate) in their experiment although videotaped interview material was provided to subjects. This effect was particularly

large when student raters participated in the study as opposed to professional recruiters who rather relied on related work experience of applicants.

Moreover, the importance of grades differs according to the occupational area examined. Certainly, job applicants in academia (PhD graduates for instance) are mainly granted access to high-quality jobs based on their previous academic achievements (Athey et al. 2007) whereas accounting graduates applying to the Big 5 accounting firms are evaluated on their GPA only in initial stages of the application process (Lewis, Shimerda and Graham 1983; Chia 2005). As a matter of fact, the relative importance of educational attainment depends on the requirements of the job (van Ours and Ridder 1991). Hence, a lack of appropriate education is one of the most frequently used immediate rejection criteria (Behrenz 2001). In addition, job-relevant education, i.e. a degree or major in the occupational field the candidate has applied for, leads to more favorable applicant ratings and hiring recommendations than educational credentials that are rather irrelevant for performing the job at stake (Knouse 1994a). McKinney et al. (2003) were also able to show that the use of GPA as a screening device depends on the individual preference of a recruiter: Some of the college recruiters in their sample seemed to rely extensively on GPA whereas others did not use grades for their selection decision at all.

A major criticism that several studies investigating the influence of grades and other academic achievement measures on application success face is the prevalent use of applications from new college graduates (Rynes, Orlitzky and Bretz 1997). As both recent graduate data and/or student evaluators are readily available to most researchers, numerous studies address success rates for entry-level jobs. But as recruiting behavior varies significantly according to the applicants' level of work experience (Rynes, Orlitzky and Bretz 1997; Cole et al. 2007), inferences from entry-level studies are likely to overestimate the power of academic qualifications. The importance of this particular kind of signal decreases over a person's career and other signals such as work experience become more decisive for an application's success (Rynes, Orlitzky and Bretz 1997; Thoms et al. 1999).

In summary, grades and other measures of educational attainment are relatively important for a candidate's application success and are one of the most important sections of a candidate's résumé (Knouse 1994a). Especially in initial screening decisions of recent graduates, GPAs are a frequently used (Cole et al. 2007), but not always the most important predictor of screening decision outcomes (McKinney et al. 2003). An exclusive use of GPAs (McKinney et al. 2003), i.e. recruiters purely relying on grades as a screening device, was found in none of the above mentioned studies.

3.3.1.2 Work Experience

A candidate's previous work experience is – similar to educational credentials – one of the items recruiters rank as crucial for selection when being asked directly for the most important selection and/or rejection criteria (Feild and Holley 1976; Lewis, Shimerda and Graham 1983; Hutchinson 1984; Pibal 1985; Harcourt and Krizan 1989; Hutchinson and Brefka 1997; Behrenz 2001; Cole et al. 2007). In Behrenz' (2001) survey for instance, a lack of appropriate experience was indicated as the most important rejection criterion. 58% of all recruiters interviewed declared that work experience is the most important characteristic in pre-selecting appropriate candidates. In final selection decisions, recruiters in the Behrenz' (2001) study do not directly mention work experience to be crucial for a positive selection outcome. Nevertheless, the attribute they consider to be decisive in interviews is something Behrenz (2001) calls professional competence which is of course increasing with job experience. Therefore, recruiters in these surveys concede the direct influential role of work experience in pre-selection and its indirect influence in employment interviews.

The preference of recruiters to use job experience as a selection device is reasonable. In contrast to educational attainment, job experience does not solely act as a signal of someone's ability or productivity, but also directly represents human capital endowments that can be of advantage in the future, i.e. for the recruiting employer. Whereas educational achievements do not directly increase an individual's productivity, but rather act as a measure of (cognitive) ability, (relevant) job experience directly leads to an accumulation of job-related knowledge which in turn increases job performance and/or supervisory ratings (Schmidt, Hunter and Outerbridge 1986). This explains the relatively high validity of previous job experience in predicting future job performance. Validity measures between 0.18 and 0.21 are reported in several studies (Hunter and Hunter 1984; Schmidt, Hunter and Outerbridge 1986; McDaniel, Schmidt and Hunter 1988). The predictive value of job experience varies according to the mean level of experience in a given occupation as well as the complexity of the job at stake: Job experience is a better predictor for low-complexity jobs (McDaniel, Schmidt and Hunter 1988) as in high-complexity jobs other factors such as educational attainment play an equally important role in predicting future performance.

As opposed to recruiter surveys, other studies rather try to reveal the 'real' importance of work experience in selection decisions and deliver more compelling evidence for actual recruiter behavior. Researchers in these studies either look at actual success rates of candidates with differing levels of job experience (Behrenz 2001; Cole et al. 2007) or

actively manipulate résumés of (fictitious) candidates with varying job experience (Hakel, Dobmeyer and Dunnette 1970; Dipboye, Arvey and Terpstra 1977; Stone and Sawatzki 1980; Fusilier and Hitt 1983; Singer and Bruhns 1991; Knouse 1994a). Job experience in these kinds of empirical studies is usually operationalized as either length of experience (measured in years), number and importance of previous jobs (e.g. assistant director of marketing), relevance of past job content (similar or dissimilar to actual job requirements) or good vs. poor employment history (length of unemployment spells, frequent quits and/or job changes etc.)

When comparing recruiters' statements to actual success rates, Behrenz (2001) and Cole et al. (2007) found that in reality, work experience does not exert as much influence on selection decisions as recruiters indicate: Behrenz (2001) found that inappropriate levels of work experience did not necessarily lead to direct elimination (in contrast to what was indicated in the survey). Comparably, Cole et al. (2007) found that work experience is not significantly related to measures of applicant employability, although the same recruiters ranked work experience as most important. Cole et al. (2007, 337) therefore conclude that "recruiters often espouse or endorse ordering of criteria as important or essential in the abstract but then utilize an alternative ordering when making actual judgments".

Nonetheless, job experience has been found to influence recruiter evaluations in several empirical investigations, especially when previous work experience is related to the job the candidate has applied for (Hakel, Dobmeyer and Dunnette 1970; Wingrove, Glendinning and Herriot 1984; Knouse 1994a). However, the magnitude of the impact work experience has on selection decisions mainly depends on the career development of the applicants examined. Whereas e.g. Cole et al. (2007) and Lewis, Shimerda and Graham (1983) find non-significant or only minor effects of work experience on recent graduates' selection success, Knouse (1994a) and Singer and Bruhns (1991) attest significant influences of previous work experience for applicants at more advanced stages of their career. Analogous to academic achievement, work experience has only been used as an independent variable in a few studies and if so, in many cases it was tested in combination with education. The only study that compared the effects of both work experience and education with each other was conducted by Singer and Bruhns (1991). They found that professional raters did not base their decision on academic qualifications at all and solely evaluated applicants based on their work experience. Student raters in this study however mainly selected applicants based on both education and experience.

Most experimental studies do not explicitly differentiate between full- or part-time work experience, but recent graduates typically possess only limited and/or irrelevant prior job experience (Nemanick and Clark 2002; Cole et al. 2007). Nevertheless, several studies report improved job-opportunities for actual graduates who have completed internships while in college (Taylor 1988; Knouse, Tanner and Harris 1999; Callanan and Benzig 2004). One of the few studies that distinctly investigated the influence of student part-time work experience in Germany was conducted by Sarcletti (2009, 2010). Investigating Bavarian graduates' time until employment after graduation, he found that internships and part-time jobs while studying reduced the time to labor market entry after graduation. But according to Sarcletti (2009), the effect is driven by increased social capital such as established contacts to potential employers rather than accumulation of human capital through job experience.

Although work experience is only seldomly studied explicitly in experiments and the investigation of actual applicant success rates, it is always among the items that recruiters rate as most important in (pre-) selection. Relevant work experience is considered less important in applications to entry-level positions, but – contrary to the influence of educational credentials – its importance increases with the career progression of applicants.

3.3.1.3 Extracurricular Activities

In addition to academic qualifications and job experience, the third major résumé content area is represented by extracurricular activities (ECAs). These activities are also commonly referred to as campus activities or simply titled applicant interests and include memberships in professional societies, college clubs, sports organizations, fraternities or sororities as well as being elected into a particular office or engaging in community activities (Cole et al. 2007). Engaging in extracurricular activities or, to be precise, mentioning these activities in an application form, is another way how candidates can differentiate from the rest of the applicant pool and signal superior abilities. As opposed to educational attainment and work experience which rather reflect cognitive abilities and/or human capital endowments, extracurricular activities may act as a signal for social competence and soft skills. Both the number and the extent of extracurricular activities may lead to a more positive employability rating, as "recruiters attribute leadership, interpersonal skill, and motivational qualities to applicants with numerous extracurricular activities" (Cole et al. 2007, 323).

As Nemanick and Clark (2002) for instance remark, the provision of extracurricular activities in résumés is of particular importance to recent graduates. College graduates'

prior work experience is typically limited or irrelevant and the academic qualifications of most applicants tend to be very similar. Therefore, providing evidence for engaging in extracurricular activities allows recruiters to differentiate among otherwise equal applicants. Lewis, Shimerda and Graham (1983) also point out that ECAs can compensate for an entry-level applicant's lack of work experience.

Only a handful of empirical studies exist that explicitly try to reveal how the provision of ECAs in résumés influences application success and affects hiring recommendations (Hakel, Dobmeyer and Dunnette 1970; Campion 1978; Lewis, Shimerda and Graham 1983; Nemanick and Clark 2002; Chia 2005; Cole et al. 2007). In survey studies, recruiters report at least some importance of ECAs in selection decisions (Lewis, Shimerda and Graham 1983; Cole et al. 2007), but the magnitude of this effect is not comparable to education and experience effects. In the Cole et al. (2007) survey, ECAs were ranked third and therefore last in importance, behind work experience and educational attainment. Lewis, Shimerda and Graham (1983) showed that in accounting settings, the importance of ECAs is dependent on the type of the employer: Whereas certified public accountants (CPA) perceive ECAs to be equally important as work experience, Fortune 500 corporations clearly prefer applicants with relevant work experience to those with extracurricular activities.

In experimental or actual success rate studies, the existence of ECAs is either measured/manipulated as the number of ECAs, the responsibility level (via an indicator for holding leadership positions) or the relevance of a certain activity to the job aimed at. Although Campion (1978) found that membership in fraternities/sororities or professional societies in combination with a good undergraduate GPA leads to most favorable recruiter evaluations more than three decades ago, little experimental work has been done on the influence of ECAs on application success. To the present, the most comprehensive lab study in this context has been conducted by Nemanick and Clark (2002). Actively manipulating (a) the number of ECAs (b) the number of leadership positions and (c) the relevance of these activities to the individual's professional career, they found various effects. They did not only report significant main effects of all three dimensions, but also showed an additional effect of number of activities and leadership positions. Raters evaluated applicants best when they showed leadership positions in many activities. Positions of leadership in a few activities or no leadership in many activities (representing a moderate level of activity) were rated second best. Holding no leadership positions in only a few activities was perceived worst. The influence of ECA-relevance for the professional career mainly depends on how actively an applicant

appears to be involved in ECAs: Whenever an applicant in the Nemanick and Clark (2002) study held leadership positions in many activities, a mixture of relevant (i.e. marketing or accounting related societies) and irrelevant (i.e. social activities) was rated best. For moderately active applicants however, business related activities were perceived as more positive than either purely social or mixed activities. Cole et al. (2007) also found support for a substantial positive relation between extracurricular activities and employability ratings. According to Chia (2005), the positive effect of extracurricular activities is only prevalent in initial stages of the selection process, i.e. in paper-based pre-selection.[26]

In Germany, only two major empirical studies investigating how ECAs might help applicants in the recruiting process exist. Although most job advertisements in Germany emphasize the importance of ECAs, Merker (2009) and Merker and Kühlmann (2010) failed to find any significant correlation between the level of ECA and (a) the number of job interview invitations, (b) time to labor market entry or (c) starting salaries of newly recruited Bavarian university graduates.[27] Similarly, Gaugler, Martin and Schneider (1995) concluded from their survey among 364 German companies that ECAs are only of minor importance in personnel selection.

All in all, recent graduates seem to be able to compensate for a lack of work experience by providing evidence for ECAs in their application (Lewis, Shimerda and Graham 1983). Nevertheless, the effect of ECAs on suitability ratings is comparably small (Hakel, Dobmeyer and Dunnette 1970) and has been revealed to be non-significant for German graduates. As only studies for entry-level positions exist so far, there is no empirical evidence on how ECAs influence selection decisions in later stages of an individual's career. Therefore, the distinctive role of ECAs needs to be studied in greater detail in the future.

[26] Although Chia (2005) explicitly addresses the possibility of indirect effects of ECA on interview success (via improved discussion and presentation skills for instance), ECAs in his study do not affect interview success. The reason might be that these 'soft skills' are already captured in the construct of Emotional Intelligence which was simultaneously tested in this study.

[27] In fact, multivariate analyses would have improved the validity of these findings.

3.3.1.4 Letters of Reference

References are reported to be a widely used selection tool by organizations all over the world. Frequently, more than 80% of organizations indicate to use any form of reference requests during their selection process (Muchinsky 1979; Gatewood and Feild 1987; Schuler, Frier and Kaufmann 1993; Dany and Torchy 1994; Schuler and Höft 2004; SHRM 2005; Cook 2009). The popularity of this selection tool is based on the same principle as information about prior work experience is: Recruiters believe that the best way to predict somebody's performance is to look at past performance. And they believe that the best way of finding out about applicants' past performance is to ask somebody who knows them well (Cook 2009). In the recruiter's belief, consulting references might deliver unique information about an applicant's past behavior that is not available from other sources (Knouse 1994b).

Typical types of references are telephone requests, standard recommendation forms or free-form letters (Cook 2009). Whereas telephone requests allow for a fast and direct communication between recommender and recruiter and are therefore very common in the US (SHRM 2004), (free-form) letters of recommendation (LORs) usually need to be interpreted by the reader alone. This type of reference is frequently used in Germany (Moser and Rhyssen 2001) and will therefore be presented in detail subsequently. Reference checks are utilized both in pre- and final selection (Moser and Cook 2009).

Unexpectedly, little empirical research dealing with LORs has been conducted so far: The study of LORs remains one of the most under-researched areas in personnel selection (Cook 2009). The existing LOR research however usually addresses one of the following three areas: (1) Why are LORs so widely used in practice? (2) How can accuracy and validity of LORs be improved? (3) How can the value of LORs to organizational decision-makers be improved?

Despite the frequent use of references in practice, academics doubt the usefulness of (free-form) reference checks as a selection tool. The average prognostic validity that LORs provide is reported to be as low as 0.13 (Aamodt, Bryan and Whitcomb 1993). Low validity measures have been reported by all studies investigating the predictive value of LORs for several success variables (e.g. Mosel and Goheen 1959; Reilly and Chao 1982; Hunter and Hunter 1984). Validity of LORs can be improved by using the keyword counting method introduced by Peres and Garcia (1962) rather than focusing on the positiveness of the letter (Aamodt, Bryan and Whitcomb 1993). But in the way LORs are frequently used in practice, they still count among the poorest predictors of future performance. Although most studies use anglo-american data, the same result has been

found in Germany as well. The only study investigating the validity of references in Germany up to now has been conducted by Moser and Rhyssen (2001). Looking at data from a German services start-up, they found the correlation between a telephone reference and the later supervisor rating to be as low as 0.20 (Moser and Rhyssen 2001). Thus, the relatively low validity has also been confirmed for the German market.

Four major problems are associated with the use of LORs that may explain the low validity of this selection tool (Aamodt, Bryan and Whitcomb 1993). These problems are (1) a letter writer's leniency, (2) limited knowledge of the applicant (and limited memory of the applicant's behavior), (3) low inter-rater reliability and (4) the influence of extraneous factors.

As illustrated in the LOR framework developed by Loher et al. (1997), applicants usually select the author of an LOR themselves. Naturally, applicants request only references from people of whom they know that they are well-disposed towards them. As a consequence, most LORs are very supportive. This is one reason for the observed leniency in LORs. Another explanation for overly positive recommendation letters are liability concerns by letter writers. Due to increased employee rights (in the US), recommendees are nowadays granted access to LORs (Aamodt, Bryan and Whitcomb 1993; Knouse 1994b). This loss of the letter content's confidentiality prevents letter writers from disclosing negative information. Guaranteed recommendee access to LORs even aggravates the leniency problem and leads to what Nicklin and Roch (2008) consider as letter inflation.

The second shortcoming of LORs is the fact that the reader usually does not know how good the letter writer knows the recommended person. Professors are for instance commonly asked to write LORs for students they have only encountered in one introductory class (Aamodt, Bryan and Whitcomb 1993). Consequently, they do not know the applicant very well. Even if the relation between recommender and recommendee is close – such as an LOR from former employers – the recommender's memory of the recommendee's behavior might be limited and/or biased. Even direct supervisors do not observe all aspects of an employee's conduct and even if they did, it is not certain that they remember all facets at the time when they are writing the LOR (Aamodt, Bryan and Whitcomb 1993).

Low inter-rater reliability is another aspect that decreases LOR credibility and consequently validity. Baxter et al. (1981) were able to show that LOR content tells more about the writer than about the recommendee. Comparing three different patterns of

agreement between 80 LORs[28], they found that two LORs written by the same recommender describing different targets (1-on-2 agreement) are more similar to each other than two LORs about the same recommendee written by different recommenders (2-on-1 agreement). They conclude that the "results are more directly a function of the writers' idiosyncrasies than of the students' qualities" (Baxter et al. 1981, 300). In a similar sense, also Judge and Higgins (1998) found that LOR content is largely driven by recommenders' affective disposition. According to their study, "positively oriented letter writers do write more favorable letters" (Judge and Higgins 1998, 217), regardless of applicant quality or suitability.

The influence of extraneous factors implies that the method used in writing an LOR seems to be more important than its content (Aamodt, Bryan and Whitcomb 1993). Also, salient factors such as recommender gender seem to have an influence on LOR content. Addressing the phenomenon of LOR writing from an evolutionary psychological perspective, Colarelli, Hechanova-Alampay and Canali (2002) were able to show that both the cooperative relationship between recommender and recommendee as well as mating interests of male recommenders positively influenced LOR favorability. Accordingly, male letter writers write more favorable letters for female applicants and LOR content is not entirely dependent on recommendee quality.

Despite the shortcomings of LORs, some studies have empirically investigated the use and perception of LORs by the reader, i.e. the recruiter. In his survey study, Behrenz (2001) found that specific LOR content is not among the most important selection or rejection criteria. Nonetheless, references are indicated to be the most important source of information by 21.5% of all interviewed recruiters. This source of information ranked third behind information gained from the personal interview and personal contacts.

By asking campus recruiters about the types of information they typically extract from LORs, Evuleocha, Ugbah and Law (2009) found that recruiters most frequently try to obtain information about somebody's ability to work with others, their work ethic, response to pressure, decision-making skills and relationship to the reference.

[28] Pattern 1 (1-on-2 agreement) measures the extent to which one perceiver, i.e. the recommender, describes two targets similarly (extent of discrimination between recommendees), pattern 2 (2-on-1 agreement) measures the extent to which two perceivers describe one target similarly (degree of consensus between different recommenders) and pattern 3 (2-on-2 agreement) measures the similarity of two perceivers' descriptions of two different targets (level of chance or stereotypic agreement). For more information on all three patterns, see Baxter et al. (1981), 296-297.

To the best of my knowledge, only few researchers (Knouse 1983; Tommasi, Williams and Nordstrom 1998; Nicklin and Roch 2008) have experimentally tested the perception of LOR content and structure. Based on his findings, Knouse (1983) advises LOR writers to give specific examples of recommendee performance as this significantly increased recruiter evaluations. The influence of negative, i.e. slightly unfavorable statements about the applicant however was not consistent and did not allow him to give advice on this aspect for LOR writers. Tommasi, Williams and Nordstrom (1998) found that most LOR readers weigh relevant LOR content, i.e. applicant qualification, more heavily than irrelevant information such as applicant and referent gender or referent status. But in line with the low inter-rater reliability problem on the writer side, they also found that there are differences in perception on the LOR reader side. That applicants benefit from inflated LORs – as opposed to non-inflated, i.e. not exaggeratedly positive, ones – was demonstrated by Nicklin and Roch (2008). Although letter readers recognize letter inflation and therefore doubt credibility of these LORs, they rate applicants having inflated LORs more positively on both hiring probability and assumed future success.

To sum up, letters of reference are a selection device frequently used by practitioners although their predictive value has been shown to be rather low. Although LORs tend to be inflated and overly supportive, the few surveys and experimental studies conducted in this area show that LOR readers are influenced by both LOR content and structure. Unfortunately, no study has tested the relative importance of LOR content in comparison to other signals or résumé characteristics so far.

3.3.1.5 Conclusion

Recapitulating what has been presented in the previous four sections, signals of ability such as academic qualification, work experience, extracurricular activities and letters of reference all represent frequently used selection devices. Each of the aforementioned ability signals has been shown to have an impact on application success. The extent to which each of these signals is able to influence the selection decision varies depending on e.g. the selection stage (pre- or final selection), the applicant's career progression and the occupational area examined.

Several authors emphasize though that there is not one single characteristic or signal which is crucial for success, but it is rather a combination of several desirable characteristics that leads to a positive selection outcome. As Cole et al. (2007, 336) put it, especially when "an applicant is above average in all résumé categories, there is no one aspect of résumé content that distinguishes, or sets apart, this applicant". The road to success in selection processes rather seems to be determined by a combination or

configuration of several of the above mentioned signals. However, only very few studies have empirically tested the simultaneous influence of multiple signals so far. Only Hakel, Dobmeyer and Dunnette (1970) and Cole et al. (2007) investigated the influence of all three résumé content dimensions simultaneously. Whereas Hakel, Dobmeyer and Dunnette (1970) reported main effects only, Cole et al. (2007) were able to show how all three dimensions also interact in the selection process. They conclude that "recruiters' perceptions of applicant employability jointly depended on the content reported in all three résumé categories" (Cole et al. 2007, 334). Not surprisingly, applicants ranked high (*low*) in all three categories received the highest (*lowest*) ratings. But applicants being perceived as weaker in one category could compensate for this deficit by being perceived as extraordinarily high in other categories. Applicants showing a high level of extracurricular activities for instance could compensate for a low work experience or weak academic performance (Cole et al. 2007).

The existence of a substitutive relationship between campus activities (i.e. ECAs) and work experience has also been assumed by Lewis, Shimerda and Graham (1983). On the other hand, van Ours and Ridder (1991) reject the hypothesis that education and experience are substitutes in the hiring process: "… education and work experience are not substituted, when hiring employees, i.e. an applicant that does not have the minimally required level of education, cannot compensate this by having more work experience (and the other way round)" (van Ours and Ridder, 1991, 218). Moser and Rhyssen (2001) conjecture that diverse signals are simultaneously utilized by recruiters. They for instance explain the popularity of LORs with their function as some kind of security which only serves to complement or confirm other application content such as work experience.

Although widely used in all stages of the selection process, the aforementioned paper-credentials seem to be most effective in pre-selection, i.e. in the decision to invite somebody to a personal interview. Furthermore, it needs to be stressed that recruiters tend to use different types of information, i.e. different signals and indices, from those they endorse when being asked directly (stated versus revealed preferences). That is the reason why quasi-experimental studies lead to superior results in this context. Especially, when it comes to more sensitive questions such as discriminatory practices in hiring, results of experimental studies should be primarily presented and discussed, as will be done in the following sections.

3.3.2 The Influence of Individual Difference Factors

A large amount of research focuses on how individual differences of all the persons involved in the selection process affect the recruitment outcome (see e.g. Posthuma, Morgeson and Campion 2002). Individual difference factors on the one hand represent applicant characteristics such as gender or race. On the other hand, some research also focuses on a variety of rater characteristics that may influence the selection outcome. The results of several empirical studies dealing with the influence of both applicant and rater characteristics during personnel selection will be presented separately.

3.3.2.1 Applicant Characteristics

In Spence's (1973) terminology individual differences among applicants are referred to as indices. Indices are unalterable characteristics of the applicant, such as gender, age, race etc. In line with Spence's (1973) theory[29], signals are assumed to have a greater impact on the employer's conditional probabilistic beliefs of the candidate's suitability than indices. Whenever indices have a major influence on the recruiter's decision although some convincing signals have been provided, this may be considered a sign of discrimination (in the workplace).[30]

As discrimination in the workplace is on the one hand directly linked to economic inefficiencies and may on the other hand evoke a sense of inequity in societies (Bendick 2007), it attracts wide interest in both the social sciences and economics. Legally, at least in industrialized countries discrimination is strictly prohibited. In the US, for instance, the *Federal Civil Rights Act* of 1964 postulates equal employment opportunities for racial/ethnic minorities, persons of non-US birth or ancestry, persons of all religions, and women. This list has been extended to also include age and disabilities by the *Age Discrimination in Employment Act* of 1967 and the *Americans with Disabilities Act* of 1990. In Europe, discrimination against all of the aforementioned characteristics plus discrimination based on an individual's sexual identity is regulated by several EU Directives[31] which have been transposed into German Law, for instance, by the commencement of the '*Allgemeines Gleichbehandlungsgesetz' (AGG)* in 2006.

---

[29] See Chapter 4.2.2 for more information about Spences (1973) theory.

[30] The only exception to this is the influence of applicant personality which will be dealt with in section 3.3.2.1.5. Although representing both an individual difference factor and an index in Spence's sense (as personality is usually unalterable), differential treatment of applicants showing varying personalities is commonly not understood as a sign of discrimination.

[31] These include the Council Directives 2000/43/EC implementing the principle of equal treatment between persons irrespective of racial or ethnic origin, 2000/78/EC establishing a general framework for equal treatment in employment and occupation and 2002/73/EC implementing the principle of equal treatment for men and women as regards access to employment, vocational training and promotion, and working conditions.

Evidence for discriminatory practices in hiring has already been found in several empirical studies dealing with diverse individual difference factors. Typically, the effect of individual difference factors is not measured in recruiter survey studies,[32] but rather through the use of (laboratory or field) experiments or actual (published) success data. Especially, correspondence and audit studies have recently been used in order to detect discrimination of minority group members. These studies were introduced as a response to the inadequacy of earlier attempts to discover discriminatory practices such as the econometric approach, i.e. regression analyses of published data (e.g. Blinder 1973; Oaxaca 1973; Ward 2001; Hinks 2002).[33] Correspondence and audit studies[34] represent carefully controlled field experiments that are able to provide the most convincing evidence on discrimination (Neumark 2011), but are also subject to substantial ethical concerns (Riach and Rich 2004; Pager 2007).

Both variants of situation testing try to reveal discrimination by using matched pairs of bogus applicants that are "identical in all relevant employment characteristics and differ only in one characteristic, such as sex, race, ethnicity or disability" (Riach and Rich 2002, F481). Audit studies are personal or in-person approaches where actual individuals – frequently professional actors – act as testers and are either sent to interviews or asked to apply on the phone. This technique enables the researcher to measure and compare the success of minority and majority applicants in both pre- and final selection, but has several shortcomings. First, coaching of testers is very costly. Second, according to critics of this approach (e.g. Heckman and Siegelman 1993), even with extensive coaching it remains impossible to control for personality differences of testers. Therefore, the fact that two applicants only differ in one characteristic and otherwise act the same cannot be guaranteed in audit studies. Third, the intransparent procedure of audit studies is criticized. The "inability to defend, or even fully enunciate, the criteria used to match

[32] Recruiters deliberately discriminating against certain minority groups usually would not reveal this behavior honestly and accurately in a survey (Riach and Rich 2002). One exception to this rule is certainly Behrenz (2001) who also asked recruiters whether or not they would reject (a) women (b) applicants over 45 years of age or (c) previously unemployed applicants. Surprisingly, 20.7% of all recruiters admitted to eliminate applicants who are 45 years or older in the first round. Still, the actual number could be substantially higher.

[33] Wage regression techniques analyze individual-level earnings of minority and majority group members and try to control for as many productivity-related factors as possible. The remaining unexplained variance is then deemed to be of discriminatory nature (Neumark 2011). As the results mainly depend on model specification, i.e. the choice of independent variables as proxies for productivity, these techniques have been heavily criticized (Riach and Rich 2002; Neumark 2011). Another inherent weakness of these techniques is that wage regression is not able to separate for instance discriminatory behavior of recruiters from 'pre-market' discrimination which is a consequence of unequal education opportunities (Bendick 2007).

[34] Outside the US, these techniques are also commonly referred to as situation testing (Bendick 2007). Other commonly used terms are employment testing, employment auditing and paired-comparison testing. All these terms will be used interchangeably throughout this thesis.

audit pair members constitutes the Achilles heel of the audit pair methodology" (Heckman and Siegelman 1993, 191).

As a remedy for most of the shortcomings of audit studies, correspondence (or résumé-based) studies have been developed and used frequently in discrimination research. Unlike in audit studies, no 'real' person is sent to an interview, but the experiment is carried out with the help of

> "carefully-matched pairs of written applications in response to advertised vacancies, to test for discrimination in labour hiring at the initial stage of selection for interview. In order to avoid detection, the letters obviously cannot be identical, but in all essential characteristics such as qualifications and experience candidates are closely matched so that the only effective distinguishing characteristic is race, ethnicity, sex, age or disability" (Riach and Rich 2002, F484).

As the researcher is truly able to control for any unintended bias through thorough matching and random assignment of variables such as letter type, this technique is less susceptible to criticism. Nevertheless, this method also has some shortcomings. First of all, the key differentiating characteristic needs to be signaled on paper (Pager 2007). While this might be conveyed easily for gender through the use of gender-specific first names, it becomes more difficult to do so with skin color for instance. Secondly, correspondence studies are only applicable to a limited range of occupations. While written applications are quite common for white-collar jobs, most blue-collar jobs (at least in the US) rather require in-person application procedures (Pager 2007). Regardless of the specific approach chosen, discrimination in these kinds of studies is usually defined as being existent "whenever two testers in a matched pair are treated differently in the aggregate or on average" (Heckman and Siegelman 1993, 198). The extent of discrimination is commonly calculated using the so called net discrimination rate "by deducting occasions of 'minority-only offered job' from occasions of 'majority-only offered job'" (Riach and Rich 2002, F491).[35]

[35] Nevertheless, not all researchers interpret their findings with the use of this net discrimination rate definition: Several studies also interpret cases in which none of the applicants was invited (or offered the job) as equal treatment and not, as in the definition of net discrimination rates, as non-observation. How this interpretation affects inferences made from correspondence and audit studies, is discussed in Riach and Rich (2002). Whenever net discrimination rates will be mentioned throughout this thesis, they have been calculated by deducting the number of cases where only the minority candidate received a callback/job offer from the number of cases where only the majority candidate received a callback/job offer divided by all cases where at least one candidate received a callback/job offer.

3.3.2.1.1 Race and Ethnicity

"Race/ethnicity/color/national origin is the personal characteristic most commonly examined" (Bendick 2007, 23) in correspondence and audit studies not only in the US, but also in Europe. Although discrimination against ethnic or racial minorities[36] is prohibited by law in all of the investigated countries, numerous studies have found evidence for such discriminatory practices in hiring.[37]

In the US, discrimination against African-American (e.g. Turner, Fix and Struyk 1991; Bendick, Jackson and Reinoso 1994; Pager 2003; Bertrand and Mullainathan 2004; Pager and Western 2005; Pager, Western and Bonikowski 2009) and Hispanic applicants (see Cross et al. 1990; Bendick et al. 1991; Firestone, Yanoff and Montenegro 2002; Pager, Western and Bonikowski 2009 for instance) has been tested and found repeatedly.[38] Discrimination does not only occur in jobs that mainly require 'only' completion of secondary school (Turner, Fix and Struyk 1991; Lodder, McFarland and White 2003), but also among university graduates (Bendick et al. 1991; Nunes and Seligman 1999).

In Britain, several ethnic minorities such as Asian, (West) Indian, Pakistani, African or Australian immigrants have been reported to have significantly lower success rates than a British-born white applicant in both initial and final stages of personnel selection processes (see e.g. Daniel 1968; McIntosh and Smith 1974; Hubbuck and Carter 1980; Esmail and Everington 1993). These findings are not limited to certain occupations, but hold true for both white-collar (e.g. Jowell and Prescott-Clarke 1970; Firth 1981; Esmail and Everington 1993) and blue-collar or unskilled occupations (see McIntosh and Smith 1974 for example). In addition, Riach and Rich (1991) were able to show that especially Vietnamese-, but also Greek-born immigrants are discriminated against in Australia.

Similarly, discrimination against immigrants has been revealed by audit or correspondence studies in other European countries such as Belgium (Smeeters and Nayer 1998), Germany (Goldberg, Mourinho and Kulke 1996), Greece (Drydakis and

[36] Racial and ethnic minorities in this thesis are understood in a broader sense and also include (first and second generation) immigrants. However, discrimination against religious minorities will not be discussed during this section.

[37] Moreover, numerous papers deal with discriminatory practices against several minorities on other markets such as housing or product markets (e.g. List 2004; Ahmed, Andersson and Hammarstedt 2009; Gneezy, List and Price 2012). However, all following sections are restricted to recruitment processes and discrimination in hiring.

[38] In all of the studies mentioned in this section, perceived race has been manipulated in correspondence studies by (a) typical and identifiable (white or minority) names or (b) indication of interests/extracurricular activities pointing to specific minority affiliation (e.g. NAACP strongly signals an African-American applicant (Pager 2007)). In-person audit studies on the other hand also made use of accents or physical appearance.

Vlassis 2010), the Netherlands (Bovenkerk et al. 1995) or Spain (de Prada et al. 1996). Here, the minority applicants were of Albanian (Drydakis and Vlassis 2010), Moroccan (Bovenkerk et al. 1995; de Prada et al. 1996; Smeeters and Nayer 1998), Turkish (Goldberg, Mourinho and Kulke 1996) or Surinamese origin (Bovenkerk et al. 1995). As it is not the purpose of this thesis to give a comprehensive review of racial or ethnic discrimination worldwide, the discussion of more detailed test results will be limited to the most recent and relevant studies.[39]

In one of the most frequently cited correspondence studies on racial discrimination, Bertrand and Mullainathan (2004) examined the effect of race-specific names on callback rates in the US. Designing résumés with either very white-sounding names (Emily Walsh and Greg Becker) or very African-American-sounding names (Lakisha Washington and Jamal Jones), they found evidence for substantial discrimination against applicants of African-American origin.[40] In their experiment, African-American applicants faced a net discrimination rate of 29.5%.[41] This racial gap remained uniform across all occupations and industries and did not vary statistically significantly with employer size. Surprisingly, even federal or equal opportunity employers did not discriminate less against African-American applicants. The only significant interaction effects Bertrand and Mullainathan (2004) report are the effects of neighborhood and résumé quality. Applicants living in a better, i.e. more affluent or 'whiter' neighborhood[42] had better chances to be called back, but this positive effect was the same for both black and white applicants. Résumé quality was (slightly) manipulated through incremental changes e.g. in labor market experience, language skills or the declaration of an email-address. In contrast to the expectations, résumé quality effects were greater for majority candidates than minority candidates, i.e. Whites benefited more from an improved résumé than black applicants did. This result prompted Bertrand and Mullainathan (2004) to reject the notion of statistical discrimination in their study as discrimination would have decreased with the use of additional ability signaling. Taste-based

[39] For a review of the most important racial discrimination studies in the US, see Bendick (2007). Riach and Rich (2002) in addition present and compare several European findings and Kolle (2012a) gives an overview of the most recent European and US studies.

[40] Bertrand and Mullainathan (2004) rejected the assumption that the name represents and conveys socio-economic background rather than racial differences by controlling for mother's education of people with all the names used in the experiment.

[41] Bertrand and Mullainathan (2004) do not explicitly state the net discrimination rate, but it can be calculated from the values they provide in table 2, p. 999: Number of white-favored cases (111) minus number of black-favored cases (46)=65 divided by the number of cases with at least one callback (1,323-1,103=220)=29.5%.

[42] Neighborhood quality was communicated through the applicant address used in the résumé.

discrimination does also not seem to be the preferred explanation of the authors.[43] Therefore, they introduce another explanation for the type of discrimination found in their study: According to Bertrand and Mullainathan (2004), employers "may use quick heuristics in reading these résumés. One such heuristic could be to simply read no further when they see an African-American name" (Bertrand and Mullainathan 2004, 1011).[44]

In a series of three correspondence tests, Swedish researchers showed that ethnic discrimination in hiring persists across all occupations, industries and skill levels in Sweden (Bursell 2007; Carlsson and Rooth 2007, 2008a; Carlsson 2010). Even more compelling, Carlsson (2010) revealed some of the factors driving discriminatory recruiter behavior (see also Carlsson and Rooth 2008a). In 2007, Carlsson and Rooth found that applicants with Middle-Eastern-sounding names such as Mohammed Ameer or Ali Said had substantially lower callback rates than applicants with a typically Swedish name (Erik Andersson for instance) in all of the occupations and skill levels they tested. According to their results, the net average discrimination rate against applicants with Middle-Eastern names was 28.9%,[45] although all recruiters that were interviewed later on indicated to treat Middle-Eastern applicants equally.[46] In addition, they were able to show that the callback gap was greater in lower-level, i.e. unskilled jobs which means that there has been less discrimination in highly skilled jobs. This result is somewhat surprising as low-level occupations are traditionally the ones with a higher share of immigrants in Sweden. On the other hand, Carlsson and Rooth (2007) assume that ability signals are more evident in high-skilled occupations which in turn reduces room for statistical discrimination. In a follow-up study, the researchers went one step further and did not only compare callback rates of natives with a Swedish name with natives with a Middle-Eastern name (also referred to as second generation immigrant), but also introduced a third fictitious applicant being a first generation immigrant with foreign qualifications. By doing so, they wanted to disentangle effects such as ethnicity, country of birth, foreign mother tongue and foreign qualifications (Carlsson 2010) and show whether it is the foreign name (indicator for preference-based discrimination) or the foreign qualification (indicator for statistical discrimination) that drives discrimination.

[43] For a more detailed discussion of the differences between statistical and taste-based discrimination, see Chapter 4.2.3.1.

[44] However, the simple notion of heuristics does not represent a viable alternative to the economic theories of discrimination. Although heuristics may be used by employers, these are certainly based on and have been put into practice either as a consequence of distaste or statistical uncertainties.

[45] Calculated from the data Carlsson and Rooth (2007) provide on page 721: (217-66)/522=0.289.

[46] Whenever at least one candidate was invited to an interview, Carlsson and Rooth (2007) addressed him or her with an interview request.

Surprisingly, they find that ethnicity per se is the main driver behind discriminatory practices of Swedish employers: Callback rates for both first and second generation immigrants are significantly lower than the ones for native applicants with a Swedish name, but are not statistically different from each other. Using an earlier version of the paper (Carlsson and Rooth 2008a), net discrimination rates can be calculated. They were 38% for the second-generation immigrant versus the native Swedish applicant and 16% for the first generation applicant versus the second generation applicant. As a result, Carlsson and Rooth (2008a) report that 77% of the total callback gap could be explained by ethnicity per se and only 23% of the differential treatment could be attributed to the foreign qualification.

Up to now, only three situation testing studies have been conducted in the German labor market all investigating discrimination against applicants with a Turkish migrational background (Goldberg, Mourinho and Kulke 1996; Kaas and Manger 2012; Kolle 2012a). In the course of several studies initiated by the ILO, Goldberg, Mourinho and Kulke (1996) conducted two studies on ethnic discrimination in Germany. The first study, a telephone audit study, revealed discrimination against second generation immigrants in semi-skilled occupations (cumulative net discrimination rate of 19%). Applicants with a Turkish-sounding name (Yilmaz Öztürk) were significantly less often invited to an interview than applicants with a typically German-sounding name (Stefan Niemeyer).[47] In a second study, Goldberg, Mourinho and Kulke (1996) tested discrimination in higher-quality occupations through a correspondence test. This time, nationality was manipulated and applicants did not only have a Turkish-sounding name, but were born in Turkey. However, this second study did not yield significant differences in callback rates of German or Turkish applicants. More recently, Kaas and Manger (2012) investigated discrimination against economics and management science students with a Turkish-sounding name applying for an internship. The only distinguishable characteristic in their correspondence study was the Turkish-sounding name of the applicant (Fatih Yilmaz or Serkan Sezer). In order to study discrimination against second- and third-generation immigrants regardless of language effects, Kaas and Manger (2012, 2) created applicants that "have German citizenship and […] were born and educated in Germany, and all of them specify 'German' as their mother tongue". Nonetheless, Kaas and Manger (2012) revealed discriminatory behavior of German internship providers: With an average net discrimination rate of 10%, applicants with a Turkish-sounding name are invited to interviews less often than their counterpart with

[47] In this study, Turkish migrational background was solely conveyed through the distinct name. Testers in the telephone study were born and raised in Germany and did not have any dialect.

a typically German name. Especially in smaller companies, having a Turkish sounding name reduces the probability of a callback significantly. However, discrimination against candidates with a Turkish name disappears with the provision of letters of references which may be seen as an indicator for the presence of statistical, rather than taste-based discrimination. Similarly, Kolle (2012a) finds a net discrimination rate of 20.5% against applicants with a Turkish-sounding name when applying for apprenticeships in Germany, but discrimination disappears with the provision of internship certificates.

To sum up, racial or ethnic discrimination persists in labor markets worldwide despite the implementation of equal opportunity rights. This discrimination is generally not limited to specific occupations or skill-levels. Net discrimination rates are overall very high, but remain generally lower in the US which might be associated with longer equal opportunity rights and greater fear of litigation for discriminatory behavior (Riach and Rich 2002). Some studies in addition have separated 'foreignness' from 'skin color' and report differential treatment according to skin color. In the US, discrimination against applicants of African-American origin seems to occur less frequently than against Hispanics. In the UK, by contrast, non-white immigrants are discriminated to a higher extent than white immigrants (Riach and Rich 2002). This emphasizes the importance of country-specific data. Racial or ethnic discrimination tends to be driven by stereotypes against certain minority groups per se rather than by fear of language problems or uncertainty about foreign qualifications. Even second and third generation immigrants born and raised in the specific country face lower callback rates. In Germany, discrimination rates in total remain comparably low which might be attributed to diverse ability signals traditionally provided in German applications. Nevertheless, ethnic discrimination also prevails in the German labor market.

3.3.2.1.2 Gender

Gender discrimination is the second most frequently investigated field in the economic discrimination literature. Mostly, female applicants are expected to be discriminated against during recruitment. A handful of audit and correspondence studies (Levinson 1975; Firth 1982; Riach and Rich 1987; Neumark et al. 1996; Nunes and Seligman 2000; Weichselbaumer 2004; Riach and Rich 2006a; Petit 2007; Carlsson and Rooth 2008b; Kolle 2012b) have been conducted both in the US and Europe for various occupations so far.[48] The most important results of these studies will be discussed briefly in this section.

Levinson (1975) was the first to conduct telephone audit tests in Atlanta (US) in either male- (or female-) dominated occupations.[49] In both fields, he found compelling evidence for a substantial discrimination against the minority, i.e. the 'sex-inappropriate' candidate: The net discrimination rate against female applicants in male-dominated occupations was reported to be 43.2%, whereas net discrimination against males in female-dominated occupations even amounted to 64.9%. In their restaurant hiring audit study, Neumark et al. (1996) hired testers to hand-deliver CVs at Philadelphia based restaurants of differing price (and therefore also pay) range. Women were revealed to have substantially lower probabilities of receiving a job offer (-40%) and being invited to an interview (-35%) in high-price restaurants. Men in turn were discriminated against by low-price restaurant owners at the final stage of the selection process and had a 40% lower probability of being offered a job in a low-price restaurant. Nunes and Seligman (2000) examined the discriminatory behavior of auto service providers (traditionally a male-dominated occupation) in the San Francisco Bay Area (US) and found a net discrimination rate against women of 27.5% for unsolicited inquiries and of 45% among those employers that actually had vacancies advertised at the time of the study. Using a natural experiment, Goldin and Rouse (2000) were able to show that the introduction of blind auditioning to hiring procedures substantially increased the chance of female musicians being hired in one of the eight major US-orchestras. From this result they inferred discriminatory recruitment practices against women in open auditions.

The first European-based correspondence study was conducted by Firth (1982) in the United Kingdom who sent written applications in response to accountant advertisements. These results show a significantly lower success rate of females in

[48] A huge body of literature also focuses on laboratory experiments conducted to test hypothetical discriminatory behavior. As laboratory results cannot reliably be transferred to real-world decision processes (Weichselbaumer 2004), only field experiments will be presented here.

[49] Male- and female-dominated occupations in audit/correspondence studies are defined by the sex-composition or segregation of this occupation by the time of investigation (see for instance Levinson (1975) or Carlsson and Rooth (2008b)).

higher status occupations such as qualified accountants or senior jobs in commerce. Using the correspondence method, Riach and Rich (1987) investigated gender discrimination in Australia for seven different occupations. They found a significant net discrimination rate against women of 11.8% for computer analyst programers and one of 16.1% for gardeners. Weichselbaumer (2004) did not only compare success rates of men and women, but also controlled for personality traits of females as a possible source of discrimination. By creating three types of job applicants – one male candidate, a masculine female and a feminine female[50] – she tested the hypothesis whether it is sex discrimination or personality traits that drive the employer's decision. Contrary to her hypothesis, no significant difference in the discrimination rates of both types of female candidates occurred: "Unfavorable treatment in masculine occupations is not significantly reduced when a woman provides a masculine identity" (Weichselbaumer 2004, 181). Both types of female candidates were equally discriminated against in the male-dominated area of network technicians (net discrimination rate of 11.8%) and men were discriminated against in the female-dominated secretary occupation (46.8%).

All four recent correspondence studies have been conducted in Europe (Riach and Rich 2006a; Petit 2007; Carlsson and Rooth 2008b; Kolle 2012b). Riach and Rich (2006a) conducted a study in segregated, male- and female-dominated occupations in England and found evidence for a substantial discrimination against men in both mixed and female-dominated occupations. The net discrimination rate against men in secretary positions amounts to 43.1% whereas discrimination against females in male-dominated jobs was found to be 'only' 23.1%. They interpret their results as being partly driven by recently implemented affirmative action policies leading to a "substantial progess in opening up professional employment opportunities to women" (Riach and Rich 2006a, 10). Petit (2007) was able to show that French employers in the financial sector mainly discriminate against young women (aged 25) applying to high-skilled administrative positions. Women aged 37 however were not discriminated against, even if they had family obligations. One possible explanation for this is taste-based co-worker discrimination – (male) workers do not accept to be supervised by young women. This in turn might decrease worker productivity causing the employer to shy away from hiring young women for high-status positions. Another explanation is statistical discrimination that induces the employer to anticipate lower long-term productivity of women due to the high probability of career interruptions as a consequence of maternal leave and family responsibilities.

[50] Gender types were conveyed to the employer with the help of résumé content that is not related to human capital, such as hobbies or the photograph (Weichselbaumer 2004, 169).

In Sweden, Carlsson and Rooth (2008b) reported only minor, i.e. one-digit net discrimination rates for the 'minority sex' in both male- and female-dominated occupations. They conclude that demand-side discrimination is not able to explain current labor market segregation in Sweden. Up to now, only one gender discrimination correspondence study has been conducted in the German labor market. Kolle (2012b) investigated the callback probabilities of female applicants in male-dominated occupations and found that the female apprenticeship candidate is overall 17% less likely to succeed in the initial selection stage. However, subdividing the sample according to the two periods the résumés were sent out, Kolle (2012b) found that discriminatory treatment only occurred in the period which was closer to the uniform commencement of apprenticeship contracts in Germany. Therefore, he assumes that short-term-hiring employers vary systematically from employers who fill in vacancies well in advance and that discriminatory behavior is directly linked to these firm characteristics.

All in all, field experiments have revealed that gender discrimination is still apparent and predominantly exists in sex-stereotyped jobs. This holds true for both men and women, i.e. it is always the 'sex-inappropriate' candidate who is discriminated against. Surprisingly, the discrimination rate against men in female-dominated occupations is consistently higher than the one against women in 'masculine' jobs. Booth and Leigh (2010) for instance find this pro-female bias especially in occupations which are heavily female-dominated (share of females greater than 80%). Nevertheless, discrimination against females seems to be most prominent in high-status positions. Two of the more recent studies (Riach and Rich 2006a; Carlsson and Rooth 2008b) however report decreasing discrimination rates of females which might be associated with the effective implementation of affirmative action policies. Nonetheless, the nature or origin of discriminatory practices – taste-based vs. statistical discrimination – cannot be revealed unequivocably by correspondence studies.

3.3.2.1.3 Age

In addition to ethnic minorities and females, another group of (potential) employees considered to be discriminated against systematically is older workers. Indirect measures of age discrimination have already hinted at a substantial disadvantage of older labor force participants. They face a higher unemployment rate, suffer from longer unemployment spells (Bendick 1983; McDonald and Chen 1994; OECD 1998) and, when re-employed, c.p. earn lower wages than their younger counterparts (Wanner and McDonald 1983). This indirect evidence however is not able to separate supply- and demand-side effects. It does not become evident from these figures whether 'older' workers withdraw from the labor market and simply do not offer their labor any longer or whether they are discriminated against by potential employers.

Direct measurement through employment testing can yield more consistent results of demand-side ageism. Age discrimination in this case means that there are fewer opportunities for older workers that cannot be attributed to lower productivity, but are only a consequence of their age (Cain 1986). Despite the growing interest in hard evidence for ageism, audit and correspondence studies in hiring have generally focused on the influence of race and gender rather than on age. The reluctance to manipulate age in these situation tests originates from several challenges this method poses when applied to applicants of different ages. As mentioned before, the inherent advantage of correspondence studies is that differential response rates to applications can be directly attributed to the one single characteristic that varies between these two applicants. Every other influence is eliminated by holding everything else constant or randomly assigning attributes and isolating these effects during the subsequent analysis. But in the case of age "there must inevitably be a variation in the job experience of the different age groups" (Riach and Rich 2006b, 2) and the ultimate principle of correspondence testing is violated. Differential treatment by employers cannot simply be explained by the variation in age, but might also be a response to differences in human capital endowments and researchers need to find ways how to match the two groups of applicants as closely as possible.[51] Another major challenge in employment testing is to select occupational fields that are theoretically suitable for applicants of all ages. According to Warr (1994), job activities can be classified into four categories according to their inherent relationship between performance and age. Age-impaired activities (category 1) are characterized by a negative correlation between age and performance. Age-counteracted (category 2) and age-neutral (category 3) activities do not show any

[51] I will address how the researchers that have already conducted age discrimination situation testing tried to handle this problem in detail when presenting their respective results.

correlation between age and performance for different reasons, whereas in age-enhanced activities (category 4) performance is positively correlated with age due to the favorable impact of experience. Researchers conducting employment testing are well advised to select category 3 activities in order to guarantee a high level of comparability between applicants of differing age groups. Additionally, occupations that are not characterized by a strong hierarchical career progression need to be investigated.

Up to now, only few employment testing studies on age discrimination exist. Bendick, Jackson and Romero (1996), Bendick, Brown and Wall (1999) and Lahey (2008) have investigated employment opportunities for older workers in the US labor market. Riach and Rich (2006b, 2007a, 2007b) have conducted three correspondence tests in France, Spain and England. The most recent studies on ageism have been conducted by Büsch, Dahl and Dittrich (2009) in Germany and Norway[52], Albert, Escot and Fernández-Cornejo (2011) in Spain and Ahmed, Andersson and Hammarstadt (2012) for the Swedish labor market.

Bendick, Jackson and Romero (1996) were the first to test age discrimination using the correspondence testing method. By sending unsolicited written applications for three different age-neutral, technical and non-technical white-collar occupations to a list of 775 employers throughout the US, they discovered an average net discrimination rate of 26.5% against the older applicant (aged 57). In order to account for the obvious job experience gap between young (32 years old) and old applicants, they provided the older applicant either with work experience unrelated to the current job (high school teacher) or indicated extended maternity leaves for women. Additionally, only credentials for the last ten years of work experience (which both of the applicants had) were handed in – a common practice in the US. Discrimination against older applicants varied significantly between geographical regions and industries. Discrimination rates were substantially higher in the South and the West (25.6% and 42.2%, respectively). Almost no discrimination was reported for the services and retail sector and older applicants had only slightly lower chances in finance, insurance and real estate companies. Employment agencies and especially manufacturing companies in contrast substantially discriminated against older applicants. The study of Bendick, Brown and Wall (1999) basically used the same framework, but was designed as an audit study and therefore allowed for additional information on age discrimination during the interview stage.

[52] This study however does not represent an audit/correspondence study in the strict sense. Also, Petit (2007) has manipulated the age of his fictitious applicants in his study, but he focused on the employment opportunities for middle-aged women only. Hence, the Petit study examines gender differences in hiring and has already been presented in section 3.3.2.1.

Again using applicants aged 32 and 57 respectively and explaining the additional 25 years of experience with irrelevant work experience, they found an overall discrimination rate of 41.2% against the older applicant. The majority of discriminatory behavior however was again found at the pre-interview stage meaning that older applicants did not even get their 'foot in the door' and were already rejected before being able to present themselves. The third US-based correspondence study has been conducted by Lahey (2008). In contrast to both previous studies, she did not compare applicants of a distinct age, but used applicants of multiple ages between 35 and 62 and subsequently categorized them into the age groups 'young' and 'old'. In order to account for the different levels of human capital as exposed through work experience, Lahey (2008) decided to use only female applicants for entry-level jobs as employers would on the one hand readily believe that a woman had taken care of her children for years. On the other hand, entry-level jobs such as cashier or secretarial work usually represent female-dominated occupations. Just as the other two US-based studies, she provided work experience credentials covering the last ten years only. Lahey (2008) found that the applicants being classified as young were 42% (46%) more likely to be invited to an interview in Massachusetts (Florida). She assumes statistical discrimination (negative stereotypes) to be the driver of discriminatory behavior as she does not find evidence for any kind of taste-based discrimination.

Three of the European correspondence tests have been conducted by Riach and Rich (2006, 2007a, 2007b). Criticizing the artificial approach of previous age discrimination studies of how to account for differences in work experience, they design the older applicant of indeed having related work experience. Rational employers should in this case prefer the older candidate (aged 47) over the younger (aged 27), less experienced one. Preferring the older applicant in this case is therefore not interpreted as discrimination against younger applicants, but simply represents rational recruitment behavior. Discrimination against older applicants will be present whenever the younger applicant is favored although the older one demonstrates substantially higher human capital (economically irrational decision). In France (2006), Spain (2007a) and England (2007b), Riach and Rich sent unsolicited written inquiries for waiter positions to restaurants throughout the country and reported average net discrimination rates against the 47-year-old, but mentally and physically active applicant of 58.1% in France, 64.5% in Spain and 28.8% in England. In France and England, discrimination was especially apparent in the capital cities – the net discrimination rates reported for London and Paris were 68.2% and 100% respectively. Similar results are reported for Swedish restaurant worker and sales assistant applicants (net discrimination rate of

61.7% against a 46-year-old candidate) by Ahemd, Andersson and Hammarstedt (2012) and 38-year-old applicants in various Spanish occupations by Albert, Escot and Fernández-Cornejo (2011). Many of these discrimination rates even exceed most discrimination rates found in race or gender studies although serving restaurant clients clearly falls into category 3 of Warr's (1994) framework and can easily be accomplished by a fit and open-minded medium-aged applicant. The most likely explanation for this employer behavior is taste-based discrimination induced by the customer, i.e. restaurant visitors want to be served by younger waiters. In England, Riach and Rich (2007b) additionally tested discrimination in two other contexts: female applicants aged 27 versus 47 in retail and recent college graduates with a general degree, e.g. in law or economics, aged 21 and 39 respectively. The 'mature age' graduate's résumé revealed that prior to entering college at the age of 35 she had been working as a secretary for eleven years and took care of her child for additional five years. Furthermore, she was designed to be divorced in order to signal low probability of future pregnancies. The interest section of her résumé did only reveal age-neutral activities. Although showing the same educational background and possessing eleven years of (somehow related) work experience, the mature graduate was heavily discriminated against (59.6% net discrimination rate). Retail sales (female clothing stores) by contrast represented the only occupation (not only in this, but also in all previous studies) where older applicants were preferred to younger candidates: Here, the significant net 'discrimination' rate against the 27-year-old applicant was 29.6% which reflects rational choices made by retail employers.

Büsch, Dahl and Dittrich (2009) adopted an approach they call 'questionnaire study' which is similar to correspondence testing, but differs in one crucial aspect: The decision whether to hire one of the fictitious applicants was not made by actual employers in the field, but by test persons (students and personnel managers). This indeed reflects only a hypothetical hiring decision, but in contrast to laboratory experiments the decision is more realistic: Neither the position nor the company in this scenario was ficticious. Other than that, all correspondence testing rules applied and age was the only variable which was manipulated systematically. The 'questionnaire study' approach allows for additional measures other than mere response rates. In a questionnaire, subjects were not only asked to select the most suitable applicant (hiring decision), but were also asked to rank the applicants on twelve capability items reflecting perceived productivity. This allowed the researchers to control for perceived productivity in age-neutral white-collar jobs. Nevertheless, the older applicant was discriminated against and had a significantly lower hiring probability of 22 percentage points in Germany and 12 percentage points in

Norway. These few correspondence/audit studies indeed reveal discriminatory practices based on applicant age in almost all labor markets. They confirm the mostly arbitrary negative stereotypes against older labor force participants (e.g. Rosen and Jerdee 1976a, 1976b; Kite et al. 2005) that other indirect measures of age discrimination[53] and laboratory experiments (e.g. Perry and Bourhis 1998; see Finkelstein, Burke and Raju 1995 or Gordon and Arvey 2004 for a meta-analytic review) have found.

3.3.2.1.4 Physical Attractiveness

"What is beautiful is good" (Dion, Berscheid and Walster 1972, 285): Beautiful people have been found inter alia to be more successful in mate selection (e.g. Adams 1977), student evaluations (Hamermesh and Parker 2003) and political elections (e.g. Adams 1977; Klein and Rosar 2005). Additionally, they are generally considered to be more intelligent (Clifford and Walster 1973; Jackson, Hunter and Hodge 1995) and socially competent (Eagly et al. 1991) than unattractive individuals. This pro-attractiveness bias is however not only limited to social life, but has also been demonstrated in the occupational domain: Even in the labor market, beauty is rewarded. By using earnings data from three major US and Canadian household surveys, Hamermesh and Biddle (1994) showed that both beauty wage premia and plainness wage penalties exist: On average, highly attractive individuals (of both genders) c.p. earn up to 10 percent more than average-looking people and individuals of below-average attractiveness earn 5 to 10 percent less than the average-looking North American.[54] Harper (2000) reports similar results for UK labor market outcomes and Mobius and Rosenblatt (2006) found attractiveness wage premia also in economic experiments.[55] Following the exclusion principle, most of them conclude that these premia and penalties must be driven by employer discrimination à la Becker (1971). In other studies, evidence for customer discrimination in Becker's sense is found: Attractive female door-to-door fundraisers are able to raise more money than their less attractive colleagues (Landry et al. 2005) and waitresses's tips were found to increase with breast size and with the mere fact of having blond hair (Lynn 2009). However, Hamermesh and Biddle (1994, 1193) also admit that "it is difficult to disentangle the effects of alternative sources of earnings differentials in the data". Endogeneity as well as unobserved variable biases are likely to occur when applying regression techniques. As it is the case with each other individual difference

[53] Self-reports of age-discriminated employees (e.g. Johnson and Neumark 1997; Purcell, Wilton and Elias 2003), employer surveys (Daniel and Heywood 2007) and macro-data (OECD 1998).

[54] Target attractiveness was measured by panel interviewers on a 5-point-scale and subsequently categorized into these three groups.

[55] Other studies that have revealed positive labor market outcomes for more attractive individuals are e.g. Biddle and Hamermesh (1998) and Parrett (2003). Hamermesh (2011) gives a comprehensive, yet rather narrative overview of how and why 'Beauty Pays'.

factor, it is impossible to include all regressors affecting an individual's wage and unambiguously attribute residual earnings differentials to attractiveness. This is the reason why again direct influence measurements such as field and laboratory experiments yield results that are superior to regression techniques. Only with the use of experiments, researchers are able to isolate the effect of beauty and hence the level of discrimination against unattractive applicants in recruitment. But manipulating attractiveness in experiments is not as straightforward as manipulating gender or race where the simple use of distinct names conveys group affiliation. According to conventional wisdom, 'Beauty is in the eye of the beholder'. In their meta-analysis, Langlois et al. (2000) however were able to reject this argument, as they report levels of agreement for adult attractiveness measures exceeding 0.9, both within and across cultures. As Hamermesh and Biddle (1994, 1176) state, "…while 'beauty is in the eye of the beholder', beholders view beauty similarly". As agreement in attractiveness evaluations is generally high, reliability of attractiveness evaluations is high and hence it is sufficient to collect and average out beauty assessments of approximately 24 raters (Henss 1992). Recently, another more objective measure of facial attractiveness has been discovered: facial (a)symmetry. Individuals possessing an ideal facial feature arrangement are considered to be more attractive and "individual attractiveness is optimized when the face's vertical distance between the eyes and the mouth is approximately 36% of its length, and the horizontal distance between the eyes is approximately 46% of the face's width" (Pallett, Link and Lee 2010, 149). These optimal proportions are also referred to as 'new' golden ratios which are interpreted "as the first 'validated' measure of facial beauty" (Lopez Bóo, Rossi and Urzúa 2012, 8).

It is worth mentioning that unattractive individuals usually do not count among the groups protected by equal opportunity legislation (Tews, Stafford and Zhu 2009). This might be an explanation for the insufficient use of field experiments and employment testing in this field – only three correspondence studies (Rooth 2009; Ruffle and Shtudiner 2011; López Bóo, Rossi and Urzúa 2012) deliberately manipulating applicant appearance exist so far. However, several laboratory experiments measuring the attractiveness bias in decision making have been conducted in organizational and social psychology. Three relatively recent meta-analyses exist (Jackson, Hunter and Hodge 1995; Langlois et al. 2000; Hosoda, Stone-Romero and Coats 2003) of which only the latter explicitly summarizes the influence of attractiveness on job-related outcomes such as selection, hiring and performance evaluation. Including 27 studies in their analysis, Hosoda, Stone-Romero and Coats (2003) report a positive relationship between attractiveness and job-related outcomes in 55 of 62 instances resulting in a weighted

mean effect size of 0.37. Additionally, potential moderators of the attractiveness bias were identified: (1) Attractiveness is equally important for men and women (regardless of job sex-type), (2) The strength of the pro-attractiveness bias is not influenced by provision of job-relevant information, (3) Compared to between-subjects designs, effect sizes are larger in within-subject designs, (4) Results do not vary significantly according to the type of rater (students vs. professionals) and hence to the type of experiment (lab vs. 'field' experiment[56]) although students are more lenient in evaluating targets, (5) The magnitude of effect sizes varies as a function of investigated outcomes and (6) Effect sizes have decreased over time.

Up to now, only two correspondence studies have tested the influence of general physical attractiveness on hiring decisions in the field (Ruffle and Shtudiner 2011; López Bóo, Rossi and Urzúa 2012). Ruffle and Shtudiner (2011) responded to job ads in Israel for positions of either high or low customer contact. They either sent applications without a picture of the male or female applicant or attached photographs of individuals previously rated as being (un)attractive by a panel of impartial judges to the résumé. A significant preference for attractive male applicants was found: compared to both men without a picture or plain-looking men, callback rates are significantly higher for attractive men. Female applicants however do not benefit from attaching attractive pictures. On the contrary, women not attaching a picture to their application are called back more frequently than both attractive and unattractive women. Albeit, this effect is not significant when recruitment is executed by employment agencies leading to the conclusion that expected contact intensity and especially female jealousy drives the beauty punishment for female applicants. López Bóo, Rossi and Urzúa (2012) however manipulated facial beauty in their correspondence test in Argentina by varying the levels of facial (a)symmetry. Applicants attaching an optimal picture, i.e. one following the golden rule of Pallett, Link and Lee (2010), receive 36% more callbacks than unattractive candidates.[57] In the 'unattractive' condition, the same person's facial symmetry was deliberately varied through the use of a computer program. As no evidence for discrimination against applicants not attaching a photograph was found, López Bóo, Rossi and Urzúa (2012) only recommend attractive applicants to attach photos to their résumé.[58]

[56] Field experiments in this context are not understood in the sense of employment testing, but simply having résumés ranked by human resource managers/recruiters instead of students.
[57] Unfortunately, no net discrimination rates were reported in these two studies.
[58] In Argentina, it is common to attach a photograph to the application, just as is the case in e.g. Germany, Sweden and other European countries.

Although most of the experimental studies report positive attractiveness biases, i.e. the more attractive person is commonly preferred, the size of the impact remains low in comparison to other information provided in the applications, such as personality and ability measures (Tews, Stafford and Zhu 2009) or educational credentials (Dipboye, Fromkin and Wiback 1975). Furthermore, pro-attractiveness biases are more subtle in nature as they differ according to target-rater-congruence in gender and attractiveness. Pro-attractiveness biases exist for opposite-sex applicants (Agthe, Spörrle and Maner 2010) and are especially pronounced for male raters evaluating female applicants (Luxen and van de Vijver 2006). This phenomenon may be interpreted as mate selection behavior from an evolutionary psychological perspective: Same-sex applicants are not systematically favored and the opposite-sex attractiveness bias only occurs when raters are told that future contact intensity between them and the applicant will be high. Especially, female (student) raters favor unattractive female applicants which can be understood as intrasexual competition (Luxen and van de Vijver 2006; Ruffle and Shtudiner 2011).[59] These perceptions of intrasexual threat however are only revealed by average looking raters (Agthe, Spörrle and Maner 2010) and did not occur consistently among professional female recruiters (Luxen and van de Vijver 2006). Marlowe, Schneider and Nelson (1996) could also show that the extent of attractiveness biases tends to decrease with the level of recruiting experience. Nonetheless, managers of all experience levels in their study were biased by attractiveness and applicant gender.

By manipulating application quality and attractiveness, Watkins and Johnston (2000) found that attractiveness is only advantageous for applicants with mediocre application quality (in terms of education, work experience and previous achievements). Applicants with a high-quality application do not benefit (additionally) from their attractiveness, while attractive applicants with a moderate application obtain almost as high judgments and hiring probabilities as attractive and highly qualified candidates. Job type in terms of expected customer (face-to-face) contact – also labeled attractiveness-relevance of the job – additionally moderates the attractiveness bias (Beehr and Gilmore 1982; Gilmore, Beehr and Love 1986; Tews, Stafford and Zhu 2009; López Bóo, Rossi and Urzúa 2012). In line with the customer discrimination notion of Becker, attractiveness is especially important in positions involving high levels of customer contact. Furthermore, Heilman and Saruwatari (1979) concluded that for female applicants also the hierarchical position of the job moderates attractiveness bias: Female attractiveness was of help when applying for nonmanagerial positions, but a hindrance in managerial positions.

[59] Given that the majority of human resources (HR) employees is female, this result might be of increased practical relevance to female applicants.

The order in which relevant information and physical appearance data are presented to raters is also able to moderate the influence of attractiveness (Cann, Siegfried and Pearce 1981). Consequently, Cann, Siegfried and Pearce (1981) advise to only allow recruiters to look at photos of applicants after having evaluated paper credentials – an idea which is implemented into anonymous applications which are very common in the US and currently tested in Europe (Krause, Rinne and Zimmermann 2010). Similarly, Tews, Stafford and Zhu (2009) recommend providing additional job-relevant information in order to minimize attractiveness effects.

Although the majority of empirical studies testing the influence of applicant appearance consider perceived facial attractiveness as a measure of beauty, an individual's personal appearance also includes other attributes such as height, weight, attire, make-up, glasses etc. The influence of applicants' weight – or more precisely obesity – on labor market outcomes has not only been tested in wage regression analyses (e.g. Cawley 2004; Conley and Glauber 2005; Fahr 2006; Han, Norton and Stearns 2009) and laboratory experiments (e.g. Pingitore et al. 1994), but also represents the first physical attractiveness characteristic actively manipulated in correspondence testing: "…there have been no previous attempts to isolate the effect of employer's perceptions of obese/unattractive job applicants on real life labor market outcomes" (Rooth 2009, 711). Rooth (2009) manipulated perceived obesity by digitally manipulating photos and measured net discrimination rates of 15.2% for male and 16.7% for female obese applicants.[60] In an attempt to reveal the drivers of discrimination, Rooth (2009) included the results of a separate attractiveness rating into the regressions and concluded that "the results for women seem to be driven by obesity, while the results for men seem to be driven by being less attractive" (Rooth 2009, 712).

Applicant attire and grooming has also been demonstrated to have an effect on hiring evaluations both for women (Mack and Rainey 1990) and for men (Kwantes et al. 2011). At least in stereotypical male occupations, traditional business attire helped applicants provide increased hiring and promotion probabilities. Altering clothing, hair, make-up and jewelry of a female applicant, Mack and Rainey (1990) reported greater hiring probabilities for well-groomed female applicants. Focusing solely on cosmetics use, Cox and Glick (1986) however found that make-up use is correlated with attractiveness, femininity and sexiness, but does not improve the level of expected job-performance. Wearing glasses however has been shown to positively bias evaluations of intelligence, diligence, honesty and trustworthiness (Manz and Lueck 1968; Argyle and McHenry

[60] Net discrimination rates calculated from Rooth (2009), table 1, 719-720.

1971; Boshier 1975; Harris, Harris and Bochner 1982), but has hitherto not been tested in (hypothetical) hiring situations. However, Brown, Henriquez and Groscup (2008) were able to show that defendants wearing eyeglasses were perceived as more intelligent and therefore judged more favorably in fictitious juror decisions.

In conclusion, personal appearance, i.e. (facial) attractiveness and other measures such as weight or attire, bias selection decisions in hiring, in particular, when expected contact intensity between rater and applicant as well as customer contact is high. Compared to other pieces of information conveyed in résumés, however, the relative impact of attractiveness is rather low. Most of the empirical, and especially experimental, research conveyed different levels of attractiveness via manipulation of photographs. Relatively few studies (e.g. Forsythe 1990; Pingitore et al. 1994) have examined the effect of applicant attractiveness – attire and obesity in the aforementioned studies – on interview instead of pre-selection outcomes.

3.3.2.1.5 Applicant Personality

Applicant personality represents one of the few individual difference factors that are not discriminatory in nature. Applicants showing certain personality traits might be favored during personnel selection processes not simply due to an interviewer's taste for or distaste against certain personalities, but because possessing these traits might be productivity enhancing for the future job. As personality traits are relatively stable over an individual's lifetime (e.g. Costa and McCrae 1997) and usually cannot be altered by the applicant, they nonetheless count among individual difference factors.

Since the mid-1980s, psychologists have agreed on a construct oriented approach for measuring and classifying an individual's personality: The Five-Factor Model (FFM) proposed by Digman (1990) and extended by Goldberg (1993) and Costa and McCrae (1992) comprises the so called Big Five personality dimensions *Openness (to Experience)*, *Conscientiousness*, *Extraversion*, *Agreeableness* and *Neuroticisim* (or labeled more positively: *Emotional Stability*).[61] Considerable consensus exists that these five major dimensions encompass the range of an individual's personality profile to a large extent and each of the dimensions is defined as follows (e.g. Tews, Stafford and Zhu 2009; see Costa and McCrae 1992 for more details): *Openness to Experience* is a personality trait

[61] Other category labels are sometimes used (e.g. *autonomy* instead of *openness to experience*), but the five dimensions have basically the same meaning and labels can be used interchangeably. Nevertheless, some researchers also measure other personality traits such as leadership skills (Kuhn and Weinberger 2005), emotional intelligence (e.g. Chia 2005), locus of control (e.g. Cook, Vance and Spector 2000; Tay, Ang & van Dyne 2006) or Type A achievement (e.g. Cook, Vance and Spector 2000) which sometimes overlap with the Big Five dimensions.

reflecting the degree to which someone appreciates adventures, unusual ideas, curiosity and the variety of experience. Individuals high on this dimension can be characterized as imaginative, original, unconventional and independent. *Conscientiousness* refers to the degree of self-discipline, sense of duty and aim for achievement. Persons possessing high levels of conscientiousness can be described as efficient, punctual, well-organized and dependable. *Extraversion* is related to an individual's preference for human contact and basically describes how outgoing people are. Highly sociable, assertive, active, energetic and talkative are adjectives commonly used to describe those high in extraversion. *Agreeableness* encompasses somebody's level of compassion and cooperation. High levels of this dimension are reflected by being altruistic, warm, generous, trusting and cooperative. *Neuroticism* (or in contrast *Emotional Stability*) refers to the tendency to easily experience unpleasant emotions. Emotionally stable people are calm, relaxed and free from worry whereas being high in neuroticism (i.e. being emotionally unstable) is expressed by high levels of anger, anxiety and depression.

Personality traits are commonly measured through self-reports of individuals. Test persons either answer several trait-related questions (e.g. "I pay attention to details" for conscientiousness) on a 5-point Likert scale which are partially positively and partially negatively coded or position themselves on a continuum between contradictory adjectives (e.g. talkative – quiet for extraversion). A frequently cited personality inventory has been developed by Costa and McCrae (1992). "Economists are only beginning to understand the relationship between personality traits and economic outcomes" (Silles 2010, 131). Only recently, economists have included non-cognitive or soft skills (i.e. personality traits) into wage regressions and have been able to explain previously unexplained variance in individuals' earnings and labor market success (e.g. Nyhus and Pons 2005; Mueller and Plug 2006; Flossmann, Piatek and Wichert 2007; Silles 2010). Additionally, numerous studies and meta-analyses (e.g. Barrick and Mount 1991; Tett, Jackson and Rothstein 1991; Salgado 1997; Hurtz and Donovan 2000) have been conducted in order to empirically test personality-performance linkages. In a second order meta-analysis, Barrick, Mount and Judge (2001) summarized the most important findings that are relatively consistent in most of the 15 studies included in their sample. In the aggregate, conscientiousness is the only valid predictor of overall job performance. Not surprisingly, well-organized, punctual and efficient individuals perform better in their jobs regardless of job type or examined performance measure. Recently, also emotional stability seems to become a more general desirable personality trait, but the predictive value of overall performance based on emotional stability is still substantially lower than the one for conscientiousness. The remaining three dimensions (agreeableness, extraversion and

openness) are only valid performance predictors for some performance measures or types of jobs. Extraversion for instance is particularly important for managerial and sales jobs involving high levels of interpersonal contact (Mount, Barrick and Stewart 1998). Emotional stability and agreeableness in turn are valid predictors only for team performance, but not for overall job performance. Following Holland's (1985) classification of occupations[62], extraversion is of particular importance in E-type jobs (Dunn et al. 1995; Cole et al. 2004) which are characterized by high levels of interpersonal contact whereas openness is particularly useful for A- and I-type jobs.

Despite this substantial evidence for the importance of at least some personality constructs in predicting job performance, only few researchers have deliberately investigated whether recruiters actually make use of these linkages. Several questions arise when examining recruiters' use of personality in selection processes: (1) Are recruiters aware of the personality-performance connection? (2) If so, how do they reveal personality traits during selection? (3) Are these perceived personality traits related to hiring decisions? (4) In which way can personality traits influence hiring decisions?

Dunn et al. (1995) have indeed shown that managers use personality descriptions – especially descriptions of applicants' conscientiousness levels – when they are readily available to them, i.e. when they are directly attached to the application. In reality however, personality profile descriptions are usually not attached to written applications nor do applicants reveal an elaborated personality profile in job interviews. Therefore, one strand of research has been dedicated to finding out whether recruiters can accurately judge the personality profile of someone they have never met before. Brown and Campion (1994) assumed that recruiters make inferences about applicants' attributes from the biodata available in résumés, but only Cole et al. (2003a, 2003b, 2004) have empirically demonstrated that specific biodata items are related to personality traits (Cole, Feild and Giles 2003a, 2003b). They could also show that recruiters reliably identify the presence of personality-related biodata-items in résumés

[62] According to a person's vocational interest and subsequent job choice, Holland (1985) classifies each individual as one of the following RIASEC-types: The **R**ealisitc (i.e. **R**-) Type has "a preference for activities that entail explicit, ordered, or systematic manipulation of objects, tools, machines, and animals" (Holland 1985, 19). The **I**nvestigative (i.e. **I**-) Type prefers "activities that entail the observational, symbolic, systematic, and creative investigation" (Holland 1985, 19) of diverse phenomena. **A**rtistic (i.e. **A**-) Types love to engage in "ambiguous, free, unsystematized activites" (Holland 1985, 20), whereas **S**ocial (**S**-) type individuals select themselves into occupations that "entail the manipulation of others to inform, train, develop, cure, or enlighten" (Holland 1985, 21). **E**nterprising (i.e. **E**-) type indivudals also prefer to manipulate others, but in order to "attain organizational goals or economic gain" (Holland 1985, 21). Finally, the **C**onventional (**C**-) Type prefers activities that "entail the explicit, ordered, systematic manipulation of data" (Holland 1985, 22).

(Cole, Feild and Giles 2003b) which could be used for inferences by recruiters. However, whether recruiters really form personality impressions from résumé data has not been tested in these studies, but e.g. in Cable and Gilovich (1998), Cole, Feild and Stafford (2005) and Cole et al. (2009). All of these come to the conclusion that recruiters indeed infer personality traits from paper credentials. These inferences however are characterized by low inter-rater reliability and low convergent validity measured as correlations between recruiters' perceived applicant personality traits and applicants' self-reported personality profiles (Cole et al. 2009).

In (simulated) interview settings, it has also been tested whether both laypersons (i.e. students) and professional observers (recruiters) accurately judge an applicant's personality. Although not being able to accurately assess all of the single dimensions, recruiters are better in holistically judging applicant personality profiles during short interview excerpts than laypersons. Student subjects however performed better in assessing single personality dimensions, but did not judge the overall profile as accurately as recruiters did (Schmid Mast et al. 2011). This is in line with the finding reported by Barrick, Patton and Haugland (2001). Recruiters' judgments correlated more strongly with self-ratings than stranger ratings, but ratings from close friends showed the highest correlation with the actual personality profile. Schmid Mast et al. (2011, 205) assume "that recruiters are better at assessing applicants as a whole … instead of assessing how applicants differ on a given personality dimension" as their job is "to recommend the best applicant as a whole and not with respect to a specific trait". Both studies emphasize that some of the personality traits could be better assessed by interviewers (openness and extraversion for instance) while other less visible, i.e. rather internal, traits such as emotional stability could not be detected by recruiters.

Regardless of inter-rater reliability and convergent validity, the question remains whether recruiters actually use their inferences in hiring or invitation decisions and whether personality really affects hiring decisions (criterion-related validity). Cole et al. (2009) found that perceived levels of extraversion, conscientiousness and openness accounted for a significant amount of variance in suitability ratings. Chia (2005) found that emotional intelligence – a trait related to agreeableness and emotional stability – positively influenced the number of subsequent job interviews and job offers for accounting graduates. Using two different occupations according to the RIASEC typology (E-type and C-type jobs), Cole et al. (2004) showed that applicants low on conscientiousness received the lowest employability ratings regardless of job type and that extraversion was only important for E-type jobs in which interpersonal skills are

highly valued. This supports the findings by Paunonen, Jackson and Oberman (1987) who reported higher employability ratings when personality congruence (i.e. perceived person-job match in terms of personality traits) was high. Caldwell and Burger (1998) pursued graduates' success in the labor market by measuring the ratio of initial and subsequent interviews as well as the number of job offers in several disciplines. Extraversion and conscientiousness were found to be positively correlated with interview success. Similar results were reported by Cook, Vance and Spector (2000) who showed that invitation to a second interview was positively correlated with achievement striving, a trait comparable to conscientiousness, and negatively correlated with trait anxiety and locus of control. Further interview success research has also confirmed that high levels of conscientiousness and extraversion are associated with interview performance (e.g. de Fruyt and Mervielde 1999; Boudreau et al. 2001). In an attempt to explain why personality traits influence hiring decisions, Tay, Ang and van Dyne (2006) introduce the interviewing self-efficacy (I-SE) as a mediator. I-SE represents the job applicants' beliefs about their interviewing capabilities which is influenced by both personal characteristics and past interview success. Another indirect influence is demonstrated by Caldwell and Burger (1998): Specific personality traits, i.e. conscientiousness and extraversion, are positively correlated with (social and background) interview preparation which in turn positively influenced interview success.

In summary, specific applicant personality traits (especially conscientiousness) are indeed positively related to overall job performance. However, recruiters experience difficulties in accurately assessing applicants' personality traits. Low values of inter-rater reliability and convergent validity are the consequence of this inability. Nevertheless, recruiters frequently rely on the personality inferences they made from both résumés and interviews (high criterion-related validity).

3.3.2.1.6 Conclusion

Individual applicant characteristics which are usually not related to productivity have repeatedly been reported to influence selection decisions. Applicants are frequently discriminated against on the basis of their race, gender, age or physical attractiveness. As most of the discrimination has been found to occur in the initial stages of a selection process, i.e. the pre-selection based on paper credentials and résumés, anonymous applications might be an appropriate remedy for discriminatory practices in the recruitment process. Particularly, applicants in countries with a long tradition of voluntarily providing plenty of irrelevant information in applications – such as Germany – might benefit from the implementation of anonymous applications.

3.3.2.2 Rater Characteristics

Just as individual differences of applicants have been shown to bias selection decisions, specific rater characteristics such as rater gender, age, race, attractiveness and rating experience or training are likely to have an impact on rater decisions. Although numerous studies include easily observable rater attributes such as gender, age or race, these are mainly treated as moderators of applicant individual difference factor effects. Rater race for instance is usually only tested in conjunction with applicant race (see e.g. Goldberg 2005; Buckley et al. 2007; McCarthy, van Iddekinge and Campion 2010) and consequently, details on the moderating effect of rater race will be presented in section 3.3.3.1 (Similarity-Attraction Effect). The same applies to rater attractiveness: Rater attractiveness is indeed able to moderate the effect of applicant attractiveness on interview outcomes (e.g. Agthe, Spörrle and Maner 2010) which has already been introduced in Chapter 3.3.2.1.4 (Applicant Attractiveness). Only for very few rater characteristics, main effects on interview outcomes have been tested so far. Of these, only rater gender and age effects will be discussed in detail in this section.[63]

Several papers explicitly deal with the differential effects of rater gender on selection decisions. Most of the empirical (field) studies find support for the hypothesis that female raters are relatively more lenient in comparison to male evaluators (e.g. Wallach and Kogan 1959; Kohn and Fiedler 1961; Warr and Knapper 1968; Deaux and Ferris 1975; London and Poplawski 1976; Muchinsky and Harris 1977; Rose and Andiappan 1978; Elliot 1981; Parsons and Liden 1984; Raza and Carpenter 1987; Andreoni and Vesterlund 2001; Chapman and Rowe 2001). According to these results, female raters tend to award generally higher applicant evaluations, regardless of applicant gender. However, this effect is only rarely reflected in final hiring recommendations. Female interviewers in Elliot's (1981) study for instance evaluated applicants' dress, person, manner, effective intelligence and disposition more favorably than did male raters. In the overall employability rating however, assessments did not differ by rater gender. These results correspond to the findings of Parsons and Liden (1984) and Raza and Carpenter (1987). Analogous to Elliot (1981), Parsons and Liden (1984) provide support for an increased leniency of female raters when evaluating nonverbal cues of applicants. Raza and Carpenter (1987) find significant positive biases in female raters' specific ratings, but not in general employability ratings. Chapman and Rowe (2001) showed that female rater generosity is especially pronounced in unstructured or semi-structured interviews. Male interviewers' ratings in their study were however not affected by

[63] For a review of other rater characteristics such as rater training and experience or rater mood effects see e.g. Posthuma, Morgeson and Campion (2002), 31-37.

interview structure. Female generosity in applicant ratings can be explained by either higher levels of altruism or social preferences for females (Andreoni and Vesterlund 2001; Croson and Gneezy 2009) or by their better ability to interpret nonverbal cues (Hall 1978) or other extraneous factors (Farina and Hagelauer 1975; Elliot 1981). Some studies investigating interviewer gender, however, were also able to show that under specific conditions, rater gender does not influence selection decisions. Abrevaya and Hamermesh (2010) did not find support for any female charity or favoritism in author-referee pairs in the economic discipline. Investigating submission success of almost 3,000 submissions to an economic journal for more than 20 years, they found that women are not more or less generous than male referees in their rejection or acceptance decisions. Although the share of females in economics is quite low – usually being indicative of same-sex favoritism – no evidence for (female) gender favoritism was found.

Rater age is also likely to affect rater evaluations. However, only very few studies have independently examined the effect of rater age on selection decisions. Ugbah and Majors (1992) recommend applicants to develop different communication strategies dependent on interviewer age as they found that younger recruiters (aged 35 or younger) perceive applicant communication behaviors differently than those aged 35 or older. In the ageism context, Finkelstein and Burke (1998) showed that older raters judged older applicants even less favorably, representing higher degrees of ageism for older raters. Quite to the contrary, Gibson, Zerbe and Franke (1993) found younger raters to rate younger workers more favorably and older raters to evaluate older workers more positively on several work-related outcomes. For performance evaluations in supervisor-subordinate relationships, Griffeth and Bedeian (1989) were able to show that younger raters gave systematically lower ratings than older raters. However, just as in a similar study on various performance measures conducted by Cleveland and Landy (1981), this effect was quite small in comparison to more relevant work-related criteria.

3.3.3 The Influence of Social Factors

In addition to ability signals and individual difference factors of both applicant and rater, social factors are also likely to have an impact on selection outcomes. Especially the employment interview is a "dynamic social process" (Swider et al. 2011, 1276). Posthuma, Morgeson and Campion (2002, 4-5) summarize:

> "At a fundamental level, the interview is a social interaction between the interviewer and the applicant. As such, a number of scholars have examined how various social factors can influence interview outcomes. This research is predicated on the notion that individuals act and reside in a social context and this context can influence their behavior and the processes and outcomes of an interview."

Of these various factors studied in this particular strand of research, only three – Applicant-Rater Similarity, Applicant Fit and Impression Management – will be discussed in the following section. Nevertheless, numerous other social factors, e.g. initial impressions, verbal and nonverbal behavior and information exchange, can affect interview outcomes.[64]

3.3.3.1 Applicant-Rater Similarity

Not only distinct applicant and rater characteristics as outlined in Chapter 3.3.2 are able to independently bias selection decisions, but especially in the interpersonal interview situation also the (mis)match between interviewer and interviewee characteristics is expected to influence selection outcomes. A similarity-attraction effect[65] is hypothesized to occur whenever candidates with similar characteristics, biographical backgrounds, attitudes or perceived personalities are unfoundedly rated more favorably by interviewers. According to the popular saying 'Birds of a feather flock together', interviewers are assumed to prefer similar applicants to dissimilar ones.[66]

Empirically, the similarity-attraction effect has not only been shown to occur in romantic (e.g. Buss 1985) and platonic personal relationships (Bahns, Pickett and Crandall 2012), but also in the organizational context: Interpersonal similarity inter alia positively affects supervisor-subordinate evaluations (Turban and Jones 1988; Vecchio and Bullis 2001), board of director selection (Westphal and Zajac 1995), venture capitalist decisions (Franke et al. 2006), buyer-seller relationships (Lichtenthal and Tellefsen 2001) and

[64] For a detailed review of the social factors not presented here, see e.g. Posthuma, Morgeson and Campion (2002, 4-14).

[65] Other commonly used terms for the same phenomenon are similar-to-me effect, similarity hypothesis, homophily principle and in-group bias. Although being derived from slightly different theoretical assumptions, these terms are used interchangeably in this thesis.

[66] See Chapter 4.2.3.2 for a theoretical derivation of this effect.

marketing channel relationships (Homburg, Schneider and Fassnacht 2002). The phenomenon is not limited to interpersonal relationships, but can also occur in interorganizational settings (e.g. Roebken 2010). But, most importantly for the present context, the similarity-attraction paradigm has also been tested empirically in both simulated (e.g. Rand and Wexley 1975; Howard and Ferris 1996; Buckley et al. 2007) and actual appointment/selection decisions (Lin, Dobbins and Farh 1992; Graves and Powell 1995, 1996; Prewett-Livingston et al. 1996; Sacco et al. 2003; McFarland et al. 2004; Goldberg 2005; Bagües and Esteve-Volart 2010; McCarthy, van Iddekinge and Campion 2010).

In the original sense, Byrne (1971) modeled attitudinal similarity to be the driver of increased interpersonal attraction (see Chapter 4.2.3.2), but as attitudinal similarity is usually unobservable and difficult to operationalize, either actual demographic similarity[67] or perceived similarity by the interviewer is used as a proxy for attitudinal similarity in empirical investigations. Applicant-rater similarity can be measured on several dimensions: either (1) purely demographic in terms of gender, race and age, (2) biographical in terms of origin and socio-economic status, (3) in terms of human capital similarity such as educational level and status, but also (work) experience and tenure, (4) relational in terms of having similar transaction partners or (5) in terms of other relevant socio-demographic dimensions such as the geographic distance between two individuals or organizations.[68]

According to the empirical results, similarity actually attracts and influences selection and appointment decisions in various occupations such as academia, college and police officer recruiting. However, the effects are generally quite small and inconsistent. Depending on the similarity dimension examined, similarity either has no effect (age similarity), positive (race similarity) or even negative repercussions (gender similarity). In none of the empirical studies, age similarity had a significant effect on rater evaluations/hiring recommendations (e.g. Lin, Dobbins and Farh 1992; Goldberg 2005).

Race similarity studies report very inconsistent results (for an overview see e.g. McCarthy, van Iddekinge and Campion 2010), which can be partly explained by their different design. Investigating the number of fouls awarded by NBA referees, Price and Wolfers (2010) found a substantially negative opposite-race effect for both black and white referees. In their field study, Lin, Dobbins and Farh (1992) investigated same-race effects in two-person panels of several racial compositions (either panels with (a) both

[67] Also labeled 'relational demography' (e.g. Goldberg 2005).
[68] However, only results for demographic similarity effects will be presented in detail.

interviewers of the same race as the applicant, (b) both different from the applicant or (c) balanced, i.e. one interviewer of the same and one of another race) during the selection for a custodial position. They found a small, but significant same-race bias for black and Hispanic applicants which was more pronounced in conventional than in situational interviews. Prewett-Livingston et al. (1996) used four-person panels of various racial compositions in the selection process of police sergeants, but were not able to include all-white or all-black panels in order to ensure fairness. As a result, they confirmed a same-race rating effect in balanced panels and a majority-race rating effect in primarily white panels, meaning that in panels with only one black and three white raters, black raters also favored the white candidate. Similarly, McFarland et al. (2004) investigated police officer applicant success as a function of the racial composition of three-person panels and reported that black raters are more prone to the similar-to-me effect, but only in predominantly black panels. Buckley et al. (2007) showed videotaped interview responses made by actual police officer applicants to several four-person panels of all possible racial compositions and found small same-race effects for both black and white raters. McCarthy, van Iddekinge and Campion (2010) reviewed highly structured interviews for entry-level managerial positions conducted by two-person panels of different race and gender. Neither gender nor race similarity affected interview outcomes in their study. Sacco et al. (2003) however were some of the few researchers to investigate the effects of racial similarity in one-on-one college recruiting interviews instead of panel interviews. As interviewers are supposed to act differently according to the interview setting (panel vs. individual)[69], they expected results different from previous studies. Analogous to McCarthy, van Iddekinge and Campion (2010), they used highly structured interviews and could not find significant similarity-attraction biases.

In terms of sex similarity biases, results are quite consistent, but in the opposite direction than expected by the similarity-attraction paradigm: Broder (1993), Graves and Powell (1995 and 1996), Goldberg (2005) and Bagües and Esteve-Volart (2010) report sex dissimilarity effects that predominantly arise from female recruiters' preference for male candidates (Graves and Powell 1995, 1996; Bagües and Esteve-Volart 2010).[70] Male recruiters were mostly not susceptible to sex (dis)similarity effects. Only Goldberg (2005), Powell and Butterfield (2002) and Walsh, Weinberg and Fairfield (1987) reported

[69] Panel interviewers might either be more or less prone to the similarity-attraction effect than interviewers in one-on-one encounters: Group dynamics à la Asch (1956) might spur stereotypic opinions, but increased accountability due to the need to present one's evaluation of the candidate might limit such biases. Furthermore, it is conceivable that similarity-attraction effects are balanced out by other panel members' ratings (see e.g. Sacco et al. 2003).

[70] This result is consistent with the increased physical attractiveness discrimination against same-sex female applicants discussed in Chapter 3.3.2.1.4.

significant sex dissimilarity effects for male recruiters who preferred female applicants (mediated by appearance). McCarthy, van Iddekinge and Campion (2010) as well as Sacco et al. (2003) and Davison and Burke (2000) failed to find any sex similarity effects when studying highly structured interviews. Only Zinovyeva and Bagues (2011) reported the theoretically anticipated same-sex favoritism for female candidates: The more females there were on academic promotion committees, the higher the chances for female applicants to be promoted to full professorship.

All in all, both actual and perceived similarity can affect organizational decisions such as interview outcomes, but the size and direction of the effect is moderated by several other factors. In line with Podolny (1994), alternative evaluation criteria such as the homophily principle become increasingly important whenever clear evaluation criteria are missing. This might explain why similarity-attraction effects are especially pronounced in academic appointment decisions where other criteria cannot always be signaled effectively and wrong choice risk is high (Roebken 2010; Fiedler and Welpe 2008; Bagües and Esteve-Volart 2010).

Other moderators of similarity-attraction effects are panel composition, interview time and structure. The composition of interview panels has been shown to moderate the effects of e.g. race- and gender-similarity effects. In short interviews, exposure to more detailed information about the applicant is low and consequently, susceptibility to homophily effects is higher (Montoya, Horton and Kirchner 2008). Highly structured interviews – if executed thoroughly – have been demonstrated to be relatively immune to similarity effects. Anticipated future interaction with the partner however is theoretically expected to positively bias similarity-attraction effects (reward-cost theory), but it has been empirically shown by Layton and Insko (1974) that the similar-to-me effect is greater when no interaction is anticipated. It is worth mentioning that significant similarity-attraction effects have been reported inconsistently for some demographic dimensions and interview situations, but even if these effects have been demonstrated, they are generally quite small (e.g. Buckley et al. 2007). In comparison to other influencing factors, similar-to-me effects are consistently smaller than e.g. measures of ability (e.g. García, Posthuma and Colella 2008) or applicant-ideal similarity (Dalessio and Imada 1984). Most of the studies investigating similarity effects investigate interview settings (final selection), only Tsai et al. (2011) have examined similar-to-me effects derived from résumés in pre-selection, but do not find any significant impact of applicant-rater similarity which they label P-P (person-person) fit.

3.3.3.2 Applicant Fit

Closely related to the previously discussed similarity-attraction effect is the notion of applicant fit. Recruiters do not only look for applicants with specific predetermined skills and abilities, but are additionally encouraged to hire applicants that (are perceived to) fit optimally into the organization. Applicant fit can hence be regarded as another factor that is able to influence hiring recommendations above and beyond other qualifications. As applicant fit is not directly observable and conveyed through other characteristics and attributes, it is frequently not measured as an independent variable, but rather as a mediator between applicant attributes and hiring recommendations (e.g. Higgins and Judge 2004; García, Posthuma and Colella 2008) or even as a dependent measure (Kristof-Brown, Barrick and Franke 2002).

Applicant fit research distinguishes between three different types of applicant fit: person-person (P-P) fit, person-job (P-J) fit and person-organization (P-O) fit, all being distinct dimensions of P-E (person-environment) fit (Jansen and Kristof-Brown 2006). As P-P fit is defined as the similarity between the individual recruiter and the applicant (Tsai et al. 2011), it has already been introduced and discussed in the previous section (similarity-attraction effects). P-J fit describes the extent to which applicants fit with the demands of the job vacancy and possess the knowledge, skills and abilities (KSAs) that are needed for executing the specific job (Kristof-Brown 2000; Tsai et al. 2011). It can be further subdivided into demands-abilities fit and needs-supplies fit (Edwards 1991). P-O fit on the other hand rather includes the extent to which the applicants' values and goals correspond to the entire organizational culture (value congruence) regardless of the specific job requirements (Cable and Judge 1997; Judge, Higgins and Cable 2000; Jansen and Kristof-Brown 2006). P-O fit can be established by selection of suitable applicants and/or by socialization within an organization (Chatman 1991). Bretz, Rynes and Gerhart (1993) as well as Kristof-Brown (2000) were able to show that P-O and P-J fit are actually discernible factors that are independently perceived and utilized by actual recruiters. Additionally, P-O fit (or firm-specific employability) constitutes a construct that has been demonstrated to be distinguishable from an applicant's general employability (Rynes and Gerhart 1990; Adkins, Russell and Werbel 1994).

Recruiters' attempts to select applicants with a substantial amount of P-O and P-J fit can be explained by empirical results demonstrating that high levels of both P-O and P-J fit are inter alia associated with lower turnover, absence rates and intentions to quit as well as higher levels of job satisfaction and organizational commitment (Chatman 1991; O'Reilly, Chatman and Caldwell 1991; Kristof 1996; Kristof-Brown, Zimmerman and

Johnson 2005). More specifically, P-O fit has been shown to be related to organization-focused outcomes such as identification and citizenship behavior, whereas P-J is rather related to job- and career-focused outcomes, e.g. career satisfaction (Cable and de Rue 2002). However, no empirical evidence for the effect of applicant fit on more objective performance measures such as productivity exists so far (Kristof 1996).

Applicant fit can either be measured directly or indirectly (Kristof 1996): Direct methods measure the perceived (i.e. subjective) fit by having recruiters rate the perceived level of applicant fit. Sample questions (see Kristof-Brown 2000 for more details) to measure perceived fit are "To what extent does this applicant fit the demands of the job" for P-J fit and "How confident are you that this applicant would be compatible with your organization" for P-O fit. Actual, i.e. objective, fit can be measured only indirectly by explicitly comparing separately obtained individual and organizational characteristics (Kristof 1996). Although being frequently inaccurate, perceived fit has a stronger influence on hiring recommendations or selection decisions (Cable and Judge 1997).

P-O and P-J fit do not only have distinct consequences within the organization, they also have different antecedents. Whereas ability signals such as work experience and education (Tsai et al. 2011) and performance expectations (García, Posthuma and Colella 2008) are positively related to P-J fit, perceived P-O is determined by work experience (Tsai et al. 2011) or values and personality traits (Kristof-Brown 2000). Nevertheless, all fit dimensions are highly intercorrelated (Kristof-Brown 2000; Tsai et al. 2011).

Especially, perceived, i.e. subjective, applicant fit is significantly and positively related to hiring recommendations in both résumé screening and interviews (Cable and Judge 1997; Kristof-Brown 2000; García, Posthuma and Colella 2008; Tsai et al. 2011). Particularly at initial stages – résumé screening and initial interviews – however, P-J fit is considerably more important than P-O fit (Bretz, Rynes and Gerhart 1993; Adkins, Russell and Werbel 1994), as the initial stages of the selection process are particularly designed to eliminate applicants who do not meet the job requirements (Kristof-Brown 2000). Value congruence in turn is not consistently related to interview invitation decisions: Some researchers do not find any correlation between P-O fit and subsequent interview invitations, but assume that P-O fit might become more important in later selection stages (e.g. Adkins, Russell and Werbel 1994). Examining actual job offer outcomes, Cable and Judge (1997) found a substantial impact of P-O fit evaluations and conclude that "work values appear to be an important element of the interviewing process" (Cable and Judge 1997, 556). Even when controlling for rater-applicant

sympathy, applicant attractiveness, work experience, GPA, sex and race, they report a 44% increase in job offer probability for a one-unit increase in P-O fit.[71]

As has been shown, (perceived) fit between the applicant and the job or the organization mediates recruiters' hiring recommendations. In particular, the employment interview is a means to establish or increase P-O fit. Additionally, fit perceptions have been shown to explain unique variance in hiring recommendations above and beyond other – more objective – selection criteria.

3.3.3.3 Self-Presentation Tactics: Impression Management

Just as in all other social interactions, "individuals will attempt to influence their exchange partner via some form of self-presentation tactic" (Swider et al. 2011, 1276) also in the employment interview. Applicants naturally engage in influence tactics, as they want to portray a suitable image and try "to present themselves in the most favorable light possible" (Swider et al. 2011, 1276). One very common way of self-presentation is known as Impression Management (IM).[72] IM tactics can be defined as attempts by interviewees to "create, maintain, protect, or otherwise alter" (Bolino et al. 2008, 1080) a desired image. At least one of these tactics is utilized by almost every applicant during an employment interview (see e.g. Stevens and Kristof 1995; Ellis et al. 2002; Levashina and Campion 2006).

IM tactics are a multifaceted phenomenon and can be classified into the following categories: (1) verbal vs. nonverbal, (2) assertive vs. defensive, (3) self-focused vs. other-focused, (4) deceptive vs. truthful, and (5) tactical vs. strategic IM tactics. Verbal IM tactics include all activities that are orally conveyed whereas nonverbal tactics reflect interviewee behaviors such as eye-contact, nodding or smiling (Kristof-Brown, Barrick and Franke 2002; Tsai, Chen and Chiu 2005; Chen, Yang and Lin 2010). Verbal tactics can be further subdivided into assertive and defensive tactics (Tedeschi and Melburg 1984). Assertive tactics describe statements in which interviewees proactively promote positive information about themselves (Proost et al. 2010), e.g. self-promotion and ingratiation tactics. These tactics are primarily used in order to create a favorable image or bolster an existing image (Swider et al. 2011). Defensive techniques such as justifications, excuses and apologies in contrast are rather reactive in nature and are intended to protect or repair a certain image (Proost et al. 2010). Depending on the focus of the conversation between applicant and interviewer, all of the verbal tactics can either

[71] It should be mentioned here that Cable and Judge (1997) report predicted probabilities derived from logit regression that should not be interpreted in a linear way.

[72] Other commonly cited dimensions of self-presentation include applicant appearance and verbal and non-verbal behavior (Swider et al. 2011).

be defined as self-focused or other-focused. Self-focused tactics "maintain attention on the candidate and allow him or her to focus the direction of the conversation in areas which will allow him or her to excel" (Kacmar, Delery and Ferris 1992, 1253). Self-promotion, i.e. actively highlighting positive traits, knowledge, skills and attributes (McFarland, Ryan and Kriska 2003; Swider et al. 2011), is the textbook example of self-focused tactics. In other-focused activities such as ingratiation[73], other-enhancement and opinion conformity however, "the applicant gives up being the focus of attention and instead employs more subtle mechanisms of influence" (Kacmar, Delery and Ferris 1992, 1253). Other-focused tactics shift the focus of the conversation to the interviewer by e.g. complimenting him or her on the interview conduct or previous achievements (Chen, Yang and Lin 2010; Proost et al. 2010).

Depending on the authencity of the message conveyed by the applicant, self-focused IM tactics can either be classified as deceptive or truthful. Whereas truthful attempts such as self-presentation are intended to manage an existing image by simply highlighting positive traits, deceptive techniques such as slight or extensive image creation are intended to create a non-existing and false image by polishing or even fabricating responses (Swider et al. 2011). Rosenfeld (1997) additionally distinguishes between tactical, i.e. focusing on bolstering a positive short-term impression only in the interview, and more long-term focused, strategic techniques aiming at conveying credibility and trustworthiness also for a future collaboration.

Self-promotion has been shown to be the most frequently used (and therefore empirically studied) self-focused and ingratiation, i.e. tactics applied in order to evoke interpersonal liking, the most commonly adopted other-focused tactic (e.g. Proost et al. 2010). Assertive tactics are employed significantly more often than defensive ones (Stevens and Kristof 1995; Kleinmann and Klehe 2010). The presence (and therefore effectiveness) of IM tactics in field settings can be measured through postinterview surveys filled out by either the interviewer or the applicant.[74] Another, more objective way to measure IM tactic use is possible whenever video- or audiotaped actual interviews are examined: An impartial coder rates the extent of certain tactics by reviewing the tapes (e.g. Stevens and Kristof 1995).

[73] Ingratiation in this context reflects all tactics "used to increase interpersonal attraction or liking by employing subtle mechanisms of influence (i.e., verbally praising the other person, conforming with the opinion of the other person)" (Proost et al. 2010, 2157).

[74] In laboratory settings, applicant IM tactics can of course be actively manipulated by the researcher.

Although the use of IM tactics is very common among all applicants, several antecedents influence the augmented use of these techniques. Certain applicant characteristics, especially personality traits, stimulate the use of specific tactics. For instance, extraverted individuals tend to engage in self-promotion activities and applicants high on agreeableness prefer the use of other-focused tactics (Kristof-Brown, Barrick and Franke 2002). But to an even larger extent, interviewer characteristics (Delery and Kacmar 1998) and situational influences such as interview format (Ellis et al. 2002; Peeters and Lievens 2006; van Iddekinge, McFarland and Raymark 2007) and IM instructions (Peeters and Lievens 2006) affect IM tactic use.

The differential effectiveness of diverse IM tactics has been empirically demonstrated in the field (Gilmore and Ferris 1989; Stevens and Kristof 1995; Delery and Kacmar 1998; Ellis et al. 2002; McFarland, Ryan and Kriska 2002; McFarland, Ryan and Kriska 2003; Higgins and Judge 2004; Tsai, Chen and Chiu 2005; Chen, Yang and Lin 2010) as well as in laboratory settings (Kacmar, Delery and Ferris 1992; Knouse 1994a; Howard and Ferris 1996; Kristof-Brown, Barrick and Franke 2002; Peeters and Lievens 2006; Varma, Toh and Pichler 2006; van Iddekinge, McFarland and Raymark 2007; Proost et al. 2010; Kleinmann and Klehe 2010; Swider et al. 2011). Self-promotion is the tactic most commonly found to be positively affecting interview outcomes such as interviewer evaluations and hiring recommendations (e.g. Stevens and Kristof 1995; Tsai, Chen and Chiu 2005). Ingratiation – the most commonly used other-focused tactic – is also positively related to interview outcomes (e.g. Proost et al. 2010), but the effect of self-focused tactics on hiring recommendations is more pronounced, i.e. self-focused tactics lead to increased hiring recommendations. Nonverbal behaviors have mainly been shown to affect interview outcomes in laboratory experiments, but only inconsistently in actual field settings (Tsai, Chen and Chiu 2005). Deceptive tactics require intensive applicant effort and concentration which on average leads to lower quality responses. Hence, deceptive image creation is likely to be noticed by interviewers and is negatively related to interview outcomes (Swider et al. 2011). But IM tactics do not only influence interview outcomes, but also pre-selection outcomes. Résumés containing IM (e.g. self-promoting descriptive or ingratiating statements) are evaluated more favorably than résumés without these statements (Knouse 1994a; Varma, Toh and Pichler 2006).

The paths through which IM tactics impinge on final rater evaluations are also quite different. Ingratiation for instance has a positive effect on interviewer affect or liking and fit, whereas self-promotion increases levels of perceived competence (Higgins and Judge 2004; Proost et al. 2010). As the positive main effect of certain IM tactics has been

demonstrated abundantly in the past decades, recent IM research focuses on the conditions under which IM tactics are most or least effective (Swider et al. 2011). Attested moderators of IM tactic effectiveness are among others interview format (McFarland, Ryan and Kriska 2003), structure (Tsai, Chen and Chiu 2005) and length (Tsai, Chen and Chiu 2005), initial interviewee impressions (Swider et al. 2011) and interviewer affectivity (Chen, Yang and Lin 2010). According to a meta-analysis conducted by Barrick, Shaffer and DeGrassi (2009), research design (field vs. laboratory setting) does not moderate the effects of IM tactics.[75] Structured interviews are likely to affect IM effectiveness: The more structured the interview, the weaker the relationship between IM tactic and evaluations (Tsai, Chen and Chiu 2005; Barrick, Shaffer and DeGrassi 2009). In addition, IM tactic use does not affect interview outcomes in role plays (McFarland, Ryan and Kriska 2003). But also IM tactic use is contingent on interview format: In situational interviews and role plays, other-focused IM tactics are predominantly used whereas self-focused IM tactics are favored by applicants being confronted with unstructured, experience based or behavior description interviews (Ellis et al. 2002; McFarland, Ryan and Kriska 2003).

Although profound understanding about the use and effectiveness of IM tactics as well as their moderators and mediators in employment interviews exists, it still remains an unresolved question whether IM tactic influence constitutes an unsolicited bias or rather a job-related and hence desirable attribute. Rosenfeld (1997, 801) postulates the latter interpretation and states that “the ability to positively ‘sell’ oneself is often a desirable attribute both in the employment interview and in later on-the-job settings”. According to Rosenfeld (1997), especially strategic IM tactics reflect desirable applicant attributes. Barrick, Shaffer and DeGrassi (2009) however only report a low predictive validity of IM tactic use. In their meta-analysis, IM tactic use was significantly correlated with interview ratings (r=.47), but only slightly correlated with job performance (r=.15).

[75] However, certain tactics have been shown to only affect outcomes in laboratory settings (e.g. nonverbal behavior).

3.3.4 The Influence of Situational and Other Extraneous Factors

Regardless of any individual peculiarities of both rater and ratee and independent from their interpersonal interaction in the selection process, extraneous factors that are not at all related to applicant performance or job requirements may as well have an impact on the selection decision. One of the most frequently researched situational influences on selection interview outcomes (and especially validity) is interview structure. Macan (2009, 204) concludes in her recent review that a "major finding in interview research a few years ago is that interviewer judgments based on structured interviews are more predictive of job performance than those from unstructured interviews". However, substantial disagreement exists among researchers about what really constitutes a (highly) structured interview (Macan 2009). According to Campion, Palmer and Campion (1997), interview structure is able to improve interview validity and reliability through enhancing either interview content or the evaluation process. They developed the following fifteen components of structure that need to be considered in order to develop truly highly structured interviews:

(1) Base Questions on a Job Analysis,
(2) Ask Exact Same Questions of Each Candidate,
(3) Limit Prompting, Follow-up Questioning, and Elaboration on Questions,
(4) Use Better Types of Questions,
(5) Use Longer Interview or Larger Number of Questions,
(6) Control Ancillary Information,
(7) Do Not Allow Questions from Candidate Until after the Interview,
(8) Rate Each Answer or Use Multiple Scales,
(9) Use Detailed Anchored Rating Scales,
(10) Take Detailed Notes,
(11) Use Multiple Interviewers,
(12) Use Same Interviewer(s) across All Candidates,
(13) Do Not Discuss Candidates or Answers between Interviews,
(14) Provide Extensive Interview Training and
(15) Use Statistical rather than Clinical Prediction.

As researchers investigating the effects of (highly) structured interviews on interview validity and reliability do not consistently incorporate all of the 15 components (see e.g. Macan 2009; McCarthy, van Iddekinge and Campion 2010), results of the moderating impact of interview structure vary substantially (Macan 2009). However, it has been

shown already in earlier chapters that increasing the level of interview structure helped e.g. reducing female rater generosity or similarity-attraction biases.

Closely linked to interview structure is the situational influence of interview panels or committees on interview outcomes. According to Campion, Palmer and Campion (1997), interviews conducted by two or more interviewers are likely to be more reliable and valid than interviews conducted by an individual evaluator alone. Empirical evidence on the impact of panel interviewing versus individual interviewing however has been contradictory and inconclusive (Dixon et al. 2002). Nevertheless, the previous discussion on similarity-attraction effects has revealed that panel composition in terms of gender or race is able to impact interview outcomes, or more specifically moderate discriminatory biases (e.g. McFarland, Ryan and Kriska 2003; Buckley et al. 2007; Bagües and Esteve-Volart 2010; Zinovyeva and Bagues 2011). In the same vein, panel size, age and functional composition can influence interview outcomes. No empirical evidence for the influence of these types of panel composition criteria on personnel selection decisions exists so far. However, Anwar, Bayer and Hjalmarsson (2012) found that trial outcomes in the US are affected by juror age: Controlling for the effect of age on jury selection, older jurors are found to be more likely to convict than their younger colleagues.

In addition to the level of interview structure, further interview design considerations (Huffcutt, van Iddekinge and Roth 2011) such as interview medium, the extent of pre-interview information given to the interviewer(s) or simply the interview time and order are extraneous factors that might – directly or indirectly – impact interview outcomes. Three main interview mediums (face-to-face, phone and videoconference) are frequently used by recruiters. However, empirical evidence of their impact on recruiter evaluations has been conflicting: Storck and Sproull (1995) for instance report that raters using videoconference technology evaluate candidates less favorably, whereas Chapman and Rowe (2001) find that applicants interviewed face-to-face are at a disadvantage. Granting interviewers access to applicant information – test scores, résumés or application blanks – prior to the interview is also likely to bias post-interview outcomes (e.g. Dipboye 1982; Phillips and Dipboye 1989; Dougherty, Turban and Callender 1994). These pre-interview impressions and subsequent interviewer expectations have been shown to result in a confirmatory bias (Huffcutt, van Iddekinge and Roth 2011) which is expressed in modified interviewer behavior.

Several empirical studies conducted in the 1970s (Hakel, Ohnesorge and Dunnette 1970; Wexley et al. 1972; Landy and Bates 1973; Heneman et al. 1975; Wexley, Sanders and Yukl 1993) have examined the influence of interviewee order effects. These studies

suggest that “an interviewer’s evaluation may be partly a function of the characteristics of preceding interviewees” (Heneman et al. 1975, 748). These contrast or assimilation effects (Landy and Bates 1973) confirm what is well-known from anecdotal evidence that raters evaluate applicants relative to other interviewees. Mediocre interviewees may then appear to be more qualified when being preceded by poorly performing candidates. However, empirical evidence for this phenomenon has mainly been found in the artificial setting of laboratory experiments and could not be detected in the field (Landy and Bates 1973). Additionally, Wexley et al. (1972) as well as Latham, Wexley and Pursell (1975) showed that an appropriate amount of training and advice can eliminate (unsolicited) contrast effects.[76]

Finally, even the time of the day an interview is scheduled is likely to influence rater decisions. Empirical evidence from judicial decisions – a field where rater objectivity is of even greater significance than in personnel selection – suggests that in each of the three daily decision sessions, the percentage of favorable rulings drastically drops at the end of each session and returns to ‘normal’ levels after the food break (Danziger, Levav and Avnaim-Pesso 2011). As this shows that even judicial decisions are not immune to extraneous factors, selection interviews presumably are not either and the scheduled interview time might exogenously predetermine part of the interviewee’s success.

[76] These findings most likely explain why no more empirical research on contrast effects has been conducted since then.

3.4 Evaluation of Relevant Literature

Recapitulating what has been discussed in the previous sections, it becomes obvious that diverse selection processes have been examined empirically so far. Numerous factors have been shown to positively or negatively affect selection outcomes, including objective ability signals, but also individual difference factors of both applicants and raters as well as social and situational factors. Stipend awarding decisions however have only been investigated explicitly by very few authors.

Middendorff, Isserstedt and Kandulla (2009) were the only researchers aiming at understanding (merit-based) stipend granting behavior in Germany. However, they only contacted successful stipend applicants, i.e. actual scholarship recipients, and compared their characteristics and attributes to those of the entire German student body. In doing so they gained in-depth information about scholarship recipients' social background for instance, but were not able to elucidate the selection process that eventually led to the uneven distribution of stipends. It is not possible to conclude from their survey whether or not students from lower social classes simply did not apply (self-selection) or were discriminated against during stipend awarding decisions. In order to learn more about potential drivers of scholarship selection decisions, empirical evidence from other selection processes was presented.

The analysis of related educational selection decisions (particularly, college admission decisions) revealed that no coherent selection criteria catalogue exists, but that the outcome is mainly dependent on the individual decision to apply to a college (self-selection) and on the institution's idiosyncratic admission policy (see e.g. Manski and Wise 1983). However, selection decisions in most of the (US-) institutions are based on both meritocratic and non-meritocratic criteria (Rigol 2003). Comparing college admission decisions to scholarship granting decisions, several similarities, but also differences can be observed. Evidently, meritocratic influence factors are likely to be identical (i.e. grades and previous academic achievement) and will most probably execute a similar influence on decisions. However, admitting somebody to college might differ from awarding somebody a stipend as the first decision is rather designed to open up the future opportunity to obtain tertiary education, i.e. providing somebody with better opportunities. Scholarship awards in turn represent a reward for previous achievements and will not necessarily be linked to somebody's future potential. Nevertheless, a lot of parallels are to be expected in these two selection decisions.

Another huge body of research which has been presented in detail is dedicated to the investigation of success factors in personnel selection. Various influence factors have been empirically examined both in the lab and in the field. Although both recruiters and scholarship raters will be susceptible to and therefore make use of similar signals and social or individual difference factors, the extent to which these factors are able to influence the final decision will vary widely.

It needs to be considered that the purpose of selection is completely different in these two situations: Whereas personnel managers search for somebody who should optimally match the job or organizational requirements and needs to execute predetermined tasks within the organization, scholarship raters want to reward somebody for previous achievement and support him or her without actually getting in touch with this person. Future contact intensity in personnel selection is usually expected to be high whereas stipend awarding evaluators are most likely to never see the applicant again after the interview. Hence, especially individual difference and social factors are expected to be less important in such situations where contact intensity is anticipated to be low – if not non-existent.

Furthermore, choosing the wrong candidate in personnel selection can become very costly for the recruiter. In a stipend awarding context on the other hand, wrong choices will not be detrimental to the individual evaluator. Consequently, evaluators are not likely to select applicants as thoroughly as recruiters do as they do not have to bear the consequences of an inefficient, i.e. suboptimal, choice. Finally, signaling one's ability might also differ according to the specific selection purpose: In personnel selection, signals are provided in order to demonstrate (potential) productivity which is not necessary in stipend awarding decisions as no 'obligation' exists to increase future productivity. Accordingly, an evaluator's utility function is likely to differ significantly from the one of a recruiter. Nevertheless, most of the above mentioned influences from either educational or personnel selection decisions serve as a suitable basis for hypotheses development in the present context.

PART B: THE DETERMINANTS OF SUCCESSFUL SCHOLARSHIP APPLICATIONS – THEORETICAL CONSIDERATIONS

4 THEORETICAL BACKGROUND

The theoretical assumptions presented in the following section begin with a sociological discussion of who forms the so-called elite in a society and therefore represents the target population of any scholarship or grant aiming at supporting future 'leaders'. Theories of how elites are formed will be elaborated. Subsequently, the decision whom to award a grant will be theoretically considered as an economic decision under uncertainty. Therefore, the basic assumptions of New Institutional Economics and the more detailed frameworks of Agency Theory and Job Market Signaling in particular will be outlined and slightly adapted to the present case. As the decision whom to award a scholarship will probably not entirely be an objective one based solely on effective signals provided by the applicants, possible evaluator biases (inefficiencies) will be explained using the neoclassical theory of discrimination as well as the similarity-attraction paradigm which is widely used in the psychological and social sciences.

4.1 SOCIOLOGICAL APPROACH: ELITES

Although the term 'elite' is currently omnipresent, a coherent definition of elites has yet to be reached in the social sciences. There are several strands of elite research that define the term 'elite' differently. These can basically be distinguished into the pre-fascist classical elite theories and the post-fascist functional elite theories. These two directions differ mainly in the understanding of how members of a certain elite are recruited from the entire population.

4.1.1 Classical Elite Theories

An aspect which most of the classical attempts to define elites have in common, is the mass-elite distinction: The 'elite' contrasts strongly with the rest of a population, which is commonly referred to as the 'crowd' or the 'masses'. That is the reason why all the classical works dealing with the notion of elites (Mosca 1896/1939; Pareto 1916/1935; Michels 1911/1970[77]) are closely linked to and subsequently have been published immediately after Le Bon's *The Crowd – A Study of the Popular Mind* (1896/2001). Although Le Bon (2001) neither uses the term 'elite' nor the one of 'ruling class', he provided the basis for the classical elite-crowd distinction by defining the (organized or

[77] In order to enable the reader to understand during which time the classical elite theories evolved, the respective publication's year of the first edition is provided in this context. The year indicated after the " / " denotes the edition which was actually used for the present thesis.

psychological) crowd as a "a single being […] subjected to the law of the mental unity of crowds" (Le Bon 2001, 13) as opposed to a "small intellectual aristocracy" (Le Bon 2001, 10) that has always created and directed civilizations.

Mosca (1939) adopts Le Bon's concept and defines the relationship between the elite on the one hand and the crowd on the other hand as follows: "In all societies [...] two classes of people appear: a class that rules and a class that is ruled. The first class, always the less numerous, performs all political functions, monopolizes power and enjoys the advantages that power brings, whereas the second, the more numerous class, is directed and controlled by the first" (Mosca 1939, 50). According to Mosca (1939), this constellation is inevitable, as the ruling class possesses mental superiority over the crowd. Affiliation to the ruling class is not directly accomplished by heredity, but rather through nurture: Certain values and traditions are passed on to the next generation. Inevitably, upper-class descendants possess these characteristics and attributes more often than a crowd offspring. Following Mosca's (1939) understanding, upward mobility (from the crowd to the ruling class) is possible whenever vertically mobile people adopt exactly these characteristics that constitute affiliation to the ruling class.

Pareto (1935) on the other hand defines elites rather functionally and counts all "people who have the highest indices in their branch of activity" (Pareto 1935, 1423 §2031), i.e. those, who are most capable in their particular field, among the class of people which is called elite. This class is further divided into a governing and a non-governing elite whereas members of the first group "directly or indirectly play some considerable part in government" (Pareto 1935, 1423 §2032) and affiliates of the latter group constitute the rest. This constellation leads to a societal trichotomy (governing elite – non-governing elite – crowd). Whether or not members of the governing elite possess superior capacity/ ability to non-governing elite members is however not clearly conveyed by Pareto (1935). According to Pareto (1935), direct and indirect heredity is a means of entering the elite class, but circulation is also essential to the continued existence of elites: Whenever elite members lose characteristics important for belonging to an elite, they descend to a lower class. Conversely, people stemming from the lower class may move up to the elite whenever they possess some crucial characteristics of this particular elite. In Pareto's (1935) sense, the use of violence and deceit is permitted – if not necessary – for a well-functioning elite: Pareto (1935) states that revolutions only emerge as a consequence of the ruling classes' reluctance to use violence.

The third fundamental classical elite theory is Michel's (1911/1970) *Law of Oligarchy*. Focusing on the possibility of intraparty political democracy, Michel (1970) posits that

each consistent organization – even a democratic political party – needs a leader. Michel (1970) identifies this fact as the beginning of the end of democracy. Conditioned by either gratitude, worship or intellectual inferiority, the crowd shows a psychological need for guidance leading to the inevitability of leadership in each form of societal life. Michel (1970) pessimistically postulates that even in democratic societies, one clique of the ruling class is simply replaced by another. Compared to aristocracy, democracy cannot be understood as a remedy for this rule, but only as the lesser of two evils.

All three classical elite theories have in common that a small number of people belonging to the elite possess materially, intellectually and psychologically superior abilities to the mentally inferior, but numerically larger crowd which is in need of guidance (Hartmann 2004). The obvious ideological proximity to fascist views has discredited all classical elite theories legitimately after World War II and led to the emergence of pluralistic functional elites which are considered to be far more important nowadays and especially better suited to match the nature and characteristics of scholarship applicants in a democratic societal system than the classical approaches.

4.1.2 Functional Elite Theories

On closer consideration, the term elite historically rather originated from the notion of functional or performance elites than from what Mosca (1939), Pareto (1935) and Michels (1970) developed: In the 18th century, the French bourgeoisie used the term in their combat against aristocracy and the clergy when fighting for equal rights and opportunities for everybody (Hartmann 2004). According to their principles, individual achievement should determine the social standing more than the social background and the circumstances somebody was born into. This notion exactly encompasses the idea of performance elites: individual achievement and performance are the primary sources of access to a certain elite. The contemporary meaning of elites is also reflected in its encyclopedic definitions: "a minority group of persons who hold positions of eminence and power in some field, especially in the social or political field" (Encyclopedia Britannica 1975, 854). German lexica define elites as follows: "[zu frz. élire »auswählen«] … eine Auslese darstellende Gruppe von Menschen mit besonderer Befähigung, besonderen Qualitäten; Führungsschicht" (Brockhaus 2006, 769)[78]. According to Meyers Grosses Taschenlexikon (2003, 1747), an elite represents a "politisch oder sozial führende Minderheit".[79]

[78] Own translation: *'[from the French language élire 'select'] … a selective group of persons with extraordinary ability or qualifications; ruling classes'.*

[79] Own translation: *'a politically or socially leading minority'.*

This definition of (functional) elites has dominated the elite discussion since the 1950s and all authors dealing with this area of research share two main assumptions (Hartmann 2004):

(1) In modern societies, there is no unique ruling class or elite, but several competing elites in different areas (functions) of societal life exist.
(2) Theoretically, access to these elites may be granted to everyone as heredity is no longer the basis for elite recruiting, but the individual's performance irrespective of the individual social background determines the affiliation to a certain elite.

The existence of plural competing functional elites goes back to Mannheim (1935/1967) who divided the formerly considered ruling class into six sub-elites: political and organizing elites, intellectual and artistic as well as moral and clerical elites. All types of elites are equally important and essential for a society, but they serve different functions: Whereas business, administration, politics and military constitute the first two types and its members are supposed to organize society, the remaining four types serve the purpose of developing a society's distinctive science and culture (Hartmann 2004).

Individual achievement being the basis for recruiting new members is the second crucial characteristic of functional elites. Heredity and/or possession recede in importance, whereas performance becomes the dominant elite selection principle for all sub-elites in industrialized societies (Dreitzel 1962; Keller 1963). Nevertheless, Dreitzel (1962) emphasizes that access to elites is only in theory open to everyone. As performance is highly correlated with education and access to education is unequally distributed among different social classes, this indirectly leads to unequal chances of success to get access to a certain elite. Dreitzel (1962) and Keller (1963) both denote that equal access to elite positions for everyone is only an ideal state that societies try to achieve. It does not necessarily mean that all elite positions are allocated solely on the basis of performance (Dreitzel 1962; Keller 1963).

As a consequence of better access opportunities for individuals from lower social classes, the magnitude of all types of elites increases and subsequently, exclusiveness decreases. The elite population hence becomes more heterogeneous which involves both risks and opportunities for the relevance of elites in a society (Hartmann 2004).

Important proponents of the currently dominating functional elite theory are e.g. Lasswell (1934, 1936), Dahrendorf (1962), Dreitzel (1962) and Keller (1963). According to Hartmann (2004), most prominent opponents of these functional elite theories are Mills

(1956, 1958) and Bourdieu (1989, 1991, 1996), but also Domhoff (1967, 1980, 1983) and Dye (1976, 1979, 1983, 1986, 1989, 1994) do not agree with the functional elite idea. They strongly disagree with all of the above mentioned authors in the following crucial aspect: They all deny the existence of several sub-elites, but again bring back the notion of one single ruling class ('Power elite' or 'Classe dominante'). Though admitting some kind of internal differentiation, they argue that this unique elite is controlled by members from the economically dominant class. This contradicts the second basic assumption of functional elite theory – openness to everyone. As social background mainly determines educational opportunities and consequently paves the way for outstanding performance, Mills (1956, 1958) and Bourdieu (1989, 1991, 1996) even deny the factual equality of access. The intensive discussion of equal opportunities on the one hand and performance as main driver of success in entering an elite on the other hand shows that there is no uniform understanding of how elites look like in industrialized societies. This might be the reason why contemporary (empirical) elite research focuses mainly on understanding the actual composition of elites and their respective power[80] rather than developing new theoretical assumptions about their structure and power.

The lack of distinct established criteria that need to be satisfied in order to rank somebody among a certain elite shows how difficult it is to identify potential elite members. As this is exactly the purpose of most merit-based scholarships, it is worthwhile to closely examine how evaluators perform such a difficult task.

[80] The most important empirical findings have been presented in Chapter 3.1.

4.2 THE ECONOMIC APPROACH

Whereas the sociological approach tries to define 'ideal' scholarship holders by identifying to which social group they will belong in the future, the economic approach basically examines the decision whom to award a stipend as an utmost rational decision. In the economic discipline, each decision can be understood as a decision under uncertainty and this also holds absolutely true in scholarship awarding decisions: The evaluator does not know the applicant, the applicants usually do not know the selection criteria and – as the previous section dealing with elites has shown – evaluators do not even exactly know whom they are looking for, but need to rank one candidate over the other. This results in several information asymmetries that have to be dealt with. How market asymmetries in general can be optimally handled, will be elaborated in the following sections.

4.2.1 Agency Theory

Apart from any sociological understanding of elites, the decision whom to award a scholarship can basically be understood as a typical principal-agent-problem (e.g. Ross 1973; Jensen and Meckling 1976; Fama 1980; Fama and Jensen 1983; Grossman and Hart 1983). Based on the ultimate assumption of New Institutional Economics – the incompleteness of information in markets and the subsequent incompleteness of contracts signed in these markets – the principal (in this case the evaluator) cannot directly observe the qualities of the agent (here: the applicant). Additionally, other components considered in principal-agent-problems are individual utility maximization of all actors and their respective risk aversion. In agency problems, three main situations may arise from the informational gaps prevailing in the market: adverse selection, moral hazard and hold-up. Whereas both moral hazard and hold-up emerge mainly from ex-post information asymmetries, adverse selection problems already exist ex ante of a contract completion.[81] Ex ante, the principal is not able to identify all characteristics of the agent who might have hidden intentions that may result in ex-post opportunistic behavior. Agents of low or below average (under the threshold level) quality will intentionally try to hide their undesirable characteristics. However, since good or above average agents are not able to distinguish themselves from the former, they might decide to exit the market. Akerlof (1970) discusses this phenomenon in his famous 'Market for Lemons' for the used car market, inter alia, and shows how hidden

[81] As the present thesis only addresses the selection of appropriate candidates and not the principal or agent behavior after the contract has been completed, only ex ante problems will be discussed in detail. For a comprehensive explanation and discussion of moral hazard and hold-up situations, see e.g. Holmström (1979,1982); Hale (2009); Picot et al. (2012).

intentions may lead to the successive market exit of above-average providers and the subsequent market breakdown. Whenever an agent with below-average qualities, i.e. an unsolicited contractual partner, is selected by the principal, the concept of adverse selection applies. The principal will only – if ever – ex post be able to detect the true characteristics and intentions of the agent.

In order to minimize or even avoid adverse selection problems, the existing information asymmetry needs to be reduced ex ante. According to agency theory principles, there are three ways how to diminish informational gaps before a contract is completed: These three are signaling, screening and self-selection which are briefly explained in the following section.

Signaling activities are initiated and conducted by the better informed party, i.e. the agent. Good or above average agents signal their ability in order to distinguish themselves from less capable, i.e. undesirable, agents. In order to do so, outstanding agents provide certain signals showing their ability. For a signal to be effective in reducing information asymmetries, two basic assumptions need to be fulfilled:

(1) For 'good' or desirable agents, the value of producing this signal needs to exceed the cost of producing it.
(2) For 'bad' or undesirable agents, the individual cost of producing the signal needs to exceed the respective value of the signal.

A frequently used example for effective signaling in hiring situations is providing university diplomas: In order to demonstrate the potential employer (principal) the future productivity (the desired attribute for the employer), able agents provide a university diploma. Regardless of any productivity gain through acquiring this diploma, this signal serves a mere allocation function. As the principal believes that the ability to acquire a university diploma is correlated with productivity, providing such a signal hence increases the value of this agent to the company. The probability of being hired by the principal increases dramatically and as a consequence thereof, the desirable agent's value of producing this signal exceeds the costs of producing it (time spent in university, waiver of income during these years = opportunity costs etc.). Undesirable agents in turn are generally not able to provide a diploma as their cost of education would exceed the value of this signal. Employers anticipate this interaction and therefore require a diploma in the present case. Spence (1973) developed a model specifically incorporating signaling activities of job market applicants and the subsequent consequences for the principal's selection process which will be presented in detail in a subsequent section.

Analogous to signaling activities accomplished by the better informed side, the principal may also engage in activities that reduce the agent's information advantage. These activities all fall into the category of screening and comprise all attempts of the principal to reveal the attributes of interest. Typical examples of screening activities in the job market include assessment centers that are designed to test certain characteristics (cognitive ability, flexibility, resistance to stress etc.) and the potential performance of applicants, but also résumé screening and contacting previous employers (reference checks) can be considered powerful screening activities.

A third way of selecting the desired agents from a pooled market with both desirable and undesirable agents is self-selection. Whenever signaling and/or screening do not lead to a satisfying result, self-selection might: Every time principals are not sure whether or not the agents incorporate the desired attributes, they might make them choose between alternative contracts. While choosing one of the contract alternatives, agents reveal some of their hitherto hidden characteristics. Prominent examples include the choice between contracts with a fixed versus a variable salary. An agent preferring the fixed rate is more likely to be a 'lazy' employee than the one volunteering to accept the variable pay. A company (principal) offering an above average wage to future employees will also be likely to demand above average effort. Following self-selection logic, only those agents who are willing (and able) to provide above average performance will apply for the job (Sadowski 2002). The same even holds true for the entire career path decision: High school graduates for instance select themselves into different areas or jobs based on what they assume about their future working conditions. Those willing to work hard and overtime in exchange for an above average salary might decide to study Business Administration or Law whereas those who value their leisure time higher than the additional earnings might start an apprenticeship in the public sector. These illustrations show how offering specific contracts may induce self-selection leading to agents sorting themselves into categories.

In the case of scholarship applications, a self-selection effect can be observed whenever the group of applicants shows significantly different, i.e. 'better', attributes than the entire student population. The principal – here the scholarship granting organization – tries to offer a contract which only attracts suited applicants. By demanding several additional documents such as letters of reference and language certificates, the organization only attracts students willing to invest enough time (and money) to be able to hand in all of these documents.

4.2.2 Spence's Job Market Signaling Theory

As already mentioned in the previous section, Spence (1973) elaborated a (principal-agent) model that explains signaling activities and their consequences in the job market. He defines hiring as an investment decision under uncertainty due to asymmetric information on job markets. Potential employers cannot observe the abilities of the unknown applicant and need to screen the applicants. Applicants need to signal their ability using a set of different signals and indices which in turn lead to employers' conditional probabilistic beliefs about the applicant's suitability. Signals are defined by Spence (1973) as alterable characteristics of an applicant and are therefore subject to manipulation by the applicant. Manipulating, i.e. improving these signals, often involves costs which Spence (1973) refers to as signaling costs. The costs to achieve a certain signal (e.g. a university degree) need to vary between two different applicants, only then a signal can be a selective signal and successfully distinguish the more suitable applicant from another less qualified one. This is known as the signaling cost condition. Indices in turn are attributes which are not alterable by the applicant such as gender, background, age, name etc. The wage offered by the employer is then a function of the applicant's signals and indices (Spence 1973).

The job market signaling model can be applied in the current situation, but some adjustments need to be made. For example, the evaluator cannot offer distinct wages or rates to differently suited applicants, but needs to decide whether or not a candidate fulfills the requirements expected in turn for a predetermined scholarship rate. Candidates will only be awarded the scholarship whenever they surpass a certain threshold, i.e. a certain expectancy level of the principal.

In line with Spence's (1973) theory, signals are assumed to have a greater impact on the employer's conditional probabilistic beliefs of the candidate's employability than indices as the former can be manipulated (improved) by the candidate and therefore follow a function of the candidate's effort (cost) and abilities. Whenever indices – which usually cannot be altered or manipulated by the applicant – have a major influence on the recruiter's decision, although some convincing signals have been provided, this may be considered a sign of discrimination (in the workplace). How an employer's evaluation may be influenced by discriminatory practices will be addressed in the following subsection.[82]

[82] The reader might also expect a discussion of Human Capital Theory (Becker 1993) in the context of investments in education. As the present thesis however deals with effective ability signaling in an academic – and not in a work-related – context, differences in human capital endowments do not immediately assist in explaining selection success in education contexts.

4.2.3 Theories Explaining Inefficiencies in Decision Making

Whenever factors that are not associated with the (potential) productivity of an individual play a major role in selection decisions, this is directly linked to a loss of economic efficiency (Bendick 2007). Both the (economic) theory of discrimination and the (sociological) similarity-attraction theory represent attempts to explain these inefficiencies in decision making.

4.2.3.1 The Theory of Discrimination

Discrimination occurs whenever members of a minority are treated less favorably than members of a majority group although they possess identical productive characteristics (Heckman 1998). Being treated less favorably can occur on several dimensions: either minority group members are offered a lower wage for the same productivity or they need to work harder (i.e. show a higher productivity) for the same wage (Arrow 1973). This is commonly referred to as income inequality or wage differentials due to discrimination in the workplace. But also in recruitment, discrimination is likely to occur and would lead to biased screening and preferential hiring of majority group members (Borjas and Goldberg 1978). The rationale behind employers' discriminatory practices may either be taste-based (Becker 1971) or statistical discrimination (Phelps 1972; Arrow 1973) besides other sociological/psychological explanations for discriminatory treatment not tackled in this work.

Taste-based discrimination implies that certain employers (or more general: decision makers) exist who have prejudices against particular groups of people. Becker (1971) calls this a certain 'taste for discrimination' which means they see an inherent disadvantage of employing minority group members. This disadvantage originates from one of the three major sources within and outside the organization: (1) Employer discrimination (2) Employee or co-worker discrimination and (3) Customer discrimination. In scenario (1), the employers' behavior lacks objectivity as they do not base their decisions solely on productivity attributes of the applicant, but express their subjective preferences when refusing an – objectively suitable – applicant. Employers that discriminate against an applicant in the second scenario do so as they are aware of their current employees' distaste for working with someone from this particular minority group. Scenario (3) explains employers' behavior by their fear to employ a person that their customers have prejudices against. Taste-based discrimination is likely to lead to income inequality as members of minority groups have to compensate for these prejudices by either accepting a lower wage for the same productivity as a majority group member or working harder (showing a higher productivity) for the same wage.

Statistical discrimination (Phelps 1972; Arrow 1973) in turn does not attribute discrimination to prejudices or tastes, but states that employers discriminate due to a lack of information about the skills and the quality of job applicants. Statistical discrimination in particular can be understood as a solution to the asymmetric information distribution. As employers only have limited information about an applicant and therefore cannot perfectly assess that person's true quality, they simply infer from prior knowledge about the group's mean performance. The decision is then based on stereotypes about the discriminated group's average (prior) performance and may lead to discriminatory hiring and/or wage differentials (Phelps 1972; Arrow 1973). For scholarship applicants belonging to certain minority groups, statistical discrimination could occur in the sense that evaluators have made unpleasant experiences (e.g. dropping out of the sponsored program) with other previously selected persons of this minority group and infer a lower quality of all applicants from their previous experience. In the present case, evaluators' previous experiences with students of the same minority group – students with a migrational background for instance – might have formed a comparably low reputation of this specific minority group in the evaluators' minds. They then – in the absence of an alternative – infer a generally lower mean performance of these students.

4.2.3.2 The Similarity-Attraction Paradigm and In-Group Favoritism

In the social sciences, it is argued that decisions of individuals are not solely driven by rational considerations. Emotional factors such as affect and interpersonal attraction are theorized to have an effect on the decision outcome as well (Berscheid and Walster 1969; Byrne 1971). In economic terms, these non-rational effects are commonly regarded as inefficiencies during the (rational) decision process as they lead to a result which is inferior to the optimal result solely based on rational considerations. One of the most frequently cited interpersonal attraction theories is the so called similarity-attraction paradigm. It hypothesizes that the (perceived) similarity between two individuals, e.g. an applicant and an evaluator, is able to influence the interpersonal attraction which in turn leads to a positive bias in the judgment of this particular person (Byrne 1971). According to Byrne and Neumann (1992), affective responses are inherent in any interpersonal encounter and do not only persist in emotionally driven decisions such as marital relationships, but also in organizational issues.

In particular, attitudinal similarity – i.e. similarity in general attitudes and values – is one of the factors that can lead to interpersonal attraction. This effect increases with the respective importance of an attitude: People are particularly attracted to others who

share similar important attitudes such as attitudes concerning war and peace or family (Byrne 1971). In the absence of any information about attitudinal similarity, also similarity in personal characteristics of two individuals may indirectly lead to positively biased judgments: Demographic similarity – e.g. in race, gender and age – causes perceived similarity of values and attitudes which then again lead to interpersonal attraction. The same holds true for biographical similarity or similarity in physical attributes such as physical attractiveness. In interview situations, both interview conduct (questioning strategy and non-verbal behavior, for instance) and information processing (i.e. what the evaluator remembers after the interview) might be affected by interpersonal attraction and lead to more favorable judgments (Byrne 1971).

Closely linked to the similarity-attraction bias are in- or intergroup biases. Intergroup bias "refers generally to the systematic tendency to evaluate one's own membership group (the in-group) or its members more favorably than a nonmembership group (the out-group) or its members" (Hewstone, Rubin and Willis 2002, 576). In the social sciences, this phenomenon can be explained by several theories, of which Social Identity Theory (Tajfel and Turner 1979) is one of the most frequently cited. In addition to merely stating that somebody is attracted to another person and consequently treats him or her favorably, social identity and other theories try to explain why in-group biases occur. Tajfel and Turner (1979) argue that in-group biases may create or reinforce a person's group identification, his or her in-group status and as a result also this person's self-esteem. Other theoretical justifications for the existence of such a similar-to-me effect include Learning Theory (Byrne 1971; Lefkowitz 2000) and Self Categorization Theory (Turner 1987; Jackson et al. 1991). But as it is not the purpose of this thesis to explain in sociological terms, why evaluators favor similar people and/or people belonging to their in-group, all of the above mentioned theories are used simultaneously to explain similarity biases. The important commonality of all these theoretical approaches is the possible positive bias towards people that are similar to the evaluator which will be analyzed in more detail in this thesis.

5 CONCEPTUAL MODEL

Before being able to empirically analyze success factors of scholarship applications, a theoretical model as well as testable hypotheses need to be developed.

5.1 THE MODEL

In a strict sense, the selection of a suitable applicant during a stipend selection process – just as in every other selection process – is preceded and determined by two independent decisions: On the one hand, the individual decision to apply for a scholarship and on the other hand the institutional or evaluator decision to award a specific applicant with the stipend. The individual decision to apply for a scholarship program (let it be A_i) is a function of the anticipated utility of applying to this program which in turn is certainly dependent on several personal characteristics of a potential applicant. However, how and why an applicant decided to send an application to this particular organization is usually unobservable by the recruiter or researcher studying selection processes. Only information about applicants who decided to apply (i.e. $A_i = 1$) is available. Therefore, the theoretical model presented below only includes the second decision, i.e. the institution's decision to award somebody a stipend, thus implicitly assuming that $A_i = 1$. This model may then be interpreted as the selection decision conditional on application.

Let us assume that evaluators during a scholarship selection process are able (or at least try) to predict the potential of a given applicant by what they can infer from the application. This generalized potential is most certainly determined by several 'sub-potentials', e.g. hard and soft skills, academic potential and non-academic potential etc. As predicted by Agency and Job Market Signaling Theory, applicants (agents) signal their ability and evaluators (principals) screen the applicants in order to reveal the characteristics and attributes of interest by interpreting both signals and indices. As has been demonstrated in the literature review, rater characteristics as well as social and situational factors might also affect the selection decision. Evaluators form an impression about each candidate by assuming his or her (unobservable) generalized potential P_i with the use of the available signals and indices. Impression formation can also be affected by all the other factors extraneous to the applicant's objective potential.

Formally, a specific applicant's potential P_i can be simply denoted as

$$P_i = \alpha_1 \overrightarrow{Sig_i} + \alpha_2 \vec{I}_i + \alpha_3 \overrightarrow{Ext_i} + \varepsilon_i$$

Note that the vector $\overrightarrow{Sig_i}$ represents all respective influence factors of applicant i which can be interpreted as ability signals in Spence's (1973) sense. These may comprise direct measures of ability such as educational attainment, e.g. grades and class rank, or experience, but also more indirect measures such as the number and intensity of extracurricular activities (ECAs), recommendations delivered in LORs, home institution characteristics (e.g. university quality and reputation), previous achievements or awards and other productivity or ability signals. The vector $\vec{I}_i$ comprises all the characteristics defined as indices by Spence (1973) that cannot be manipulated by the applicant but which are nevertheless revealed in an application. Examples of these indices are all applicant individual difference factors such as gender, race or age. $\overrightarrow{Ext_i}$ stands for all other (extraneous) factors that might affect impression formation and decision making such as rater characteristics or social and situational influences. These are most likely to not only have a direct influence on the perceived potential P_i, but are also able to moderate the influence of both signals and indices. That is the reason why the model describing applicant potential needs to be amended by two interaction terms:

$$P_i = \alpha_1\overrightarrow{Sig_i} + \alpha_2\vec{I}_i + \alpha_3\overrightarrow{Ext_i} + \alpha_4\overrightarrow{Ext}_i\,\overrightarrow{Sig_i} + \alpha_5\overrightarrow{Ext_i}\vec{I}_i + \varepsilon_i$$

Similarly, different applicant signals and indices are not assumed to be substitutes, but will presumably have an additive, i.e. a complimentary effect, on perceived applicant potential. This interaction takes into account that certain ability signals and demographic characteristics (indices) might reinforce themselves:

$$P_i = \alpha_1\overrightarrow{Sig_i} + \alpha_2\vec{I}_i + \alpha_3\overrightarrow{Ext_i} + \alpha_4\overrightarrow{Ext_i}\,\overrightarrow{Sig_i} + \alpha_5\overrightarrow{Ext_i}\vec{I}_i + \alpha_6\overrightarrow{Sig_i}\,\vec{I}_i + \varepsilon_i$$

Note that the disturbance term ε_i contains all of the other attributes that might influence an evaluator's perception of an applicant's potential which we are, however, not able to measure or observe. This can e.g. include the more subjective impression evaluators are able to form during an interview.

As the maximum number of stipends awarded in a given selection period is usually limited (as a consequence of certain budget constraints), an applicant's success is not only determined by his or her individual potential, but has to be evaluated in relation to the competing applicants' potential. In contrast to labor market and especially recruitment decisions, evaluators are not able to offer distinct wages as a function of an individual applicant's potential. The scholarship rate is predetermined by organization j, and therefore successful candidates need to surpass a certain threshold potential level, called L_j, in order to be awarded the stipend. This threshold level is in turn determined by both minimum requirements of organization j (O_j) and the applicant pool's average potential ($\overline{P_i}$):

$$L_j = \beta_1 O_j + \beta_2 \bar{P}_i + \omega_j$$

Again, we allow for some noise in the formation of the threshold level L_j by including the disturbance term ω_j. As a consequence of a minimum applicant 'quality', the unobservable selection probability of applicant i (S_i) depends on both the applicant's individual potential P_i and the threshold level L_j. More specifically, in order to be awarded the stipend, the individual potential needs to exceed the threshold level:

$$\Pr(P_i - L_j > 0) = S_i$$

$$= \Pr\Big(\big(\alpha_1 \overrightarrow{Sig_i} + \alpha_2 \vec{I}_i + \alpha_3 \overrightarrow{Ext_i} + \alpha_4 \overrightarrow{Ext_i}\,\overrightarrow{Sig_i} + \alpha_5 \overrightarrow{Ext_i}\vec{I}_i + \alpha_6 \overrightarrow{Sig_i}\,\vec{I}_i + \varepsilon_i\big) - \big(\beta_1 O_j + \beta_2 \overline{P_i} + \omega_j\big) > 0\Big)$$

5.2 Hypotheses Development

In concordance with functional elite theory, an individual's performance represents the only determinant of elite affiliation. Consequently, evaluators awarding stipends aimed at supporting future elite members, individual (past) performance should determine selection success to a large extent. Past performance however is reflected in certain ability signals applicants (agents) provide in their application. During a (perfectly) rational decision process, evaluators (principals) should base their decision solely on these credible signals of ability in order to evaluate a candidate's past performance and future potential. Consequently, Hypothesis 1 states as follows

Hypothesis 1: The provision of credible ability signals increases an applicant's chances to be awarded the stipend.

Ability signals provided in a scholarship application include various aspects of an individual's academic and non-academic performance. These include direct proofs of academic performance such as school and university grade certificates which applicants are usually required to hand in with an application. But also more indirect information about an individual's (previous) performance such as university quality or reputation, work experience (internships), extracurricular activities and letters of recommendation are usually available in written applications. Additionally, any further credential candidates provide in their application may (even inadvertently) serve as a signal of their suitability. Additional certificates and prior awards, but also remarkable rhetoric skills for instance may also signal an increased suitability for being awarded a stipend. However, evaluators need to be able to interpret all signals in order to decide whether or not they are both credible and expedient for stipend awarding. The interpretation of certain signals is facilitated whenever both their direction and range is well-known to all evaluators. High school and university grades for instance represent signals evaluators (usually professors or academic assistants) are familiar with. Furthermore, grades (at least within one country or educational system) are standardized and considered to be an objective measure of performance. As a consequence, the amount of interpretation needed is minimal and grades are likely to represent the most credible and reliable ability signal provided in an application. Hence, Hypothesis 2 can be derived as follows

Hypothesis 2: Among all ability signals, educational attainment signals (especially grades) will have the strongest effect on the probability of being awarded the stipend.

Just as in other more extensively researched selection processes with a huge supply-demand imbalance, the scholarship selection process will most likely be divided into several stages. The pre-selection (or sifting) stage of any selection process is aimed at identifying appropriate and inappropriate candidates based on their résumés and application forms. In personnel selection, appropriateness in this sense can be defined as fulfilling all the necessary requirements (i.e. KSAs) for executing this particular task and this can be captured in the applicant P-J fit construct. As the empirical literature on applicant fit has shown, P-J fit is mostly evaluated in pre-selection and only candidates who fulfill (most of) the job requirements (high P-J fit) will be further considered in the selection process and invited to job interviews for instance. In subsequent stages, recruiters then compare applicants who are theoretically (i.e. on paper) all able to perform the task. In personnel recruitment, P-O fit assessments, e.g. value congruence between organization and applicant, become more important in these stages as P-J fit has usually already been determined at an earlier stage. In the stipend awarding context, something like P-J fit might also exist, representing the individual ability to perform the 'task' the award is linked to, which is usually 'studying' in the context of university stipends. P-J fit (or its equivalent in the scholarship awarding context) is assessed during pre-selection based on the ability signals provided in an application and inappropriate candidates will already be eliminated during this early stage. During final selection (in an interview for instance) however, only applicants are considered who have been found to fulfill the minimum requirements (acceptable P-J fit) and variance in P-J fit (and therefore in the quality of ability signals) will be substantially lower in final selection. Therefore, Hypothesis 3 is developed.

Hypothesis 3: The influence of ability signals will be more important in paper-based pre-selection than in (person-to-person) final selection.

In contrast to personnel selection, evaluators in a stipend awarding process do not necessarily have an ideal candidate in mind as no direct task exists which needs to be fulfilled by the successful applicant. Consequently, no consensus exists among evaluators concerning the explicit characteristics of a stipend awardee. On the contrary, evaluators look for applicants promising to become future elite members which can be defined by diverse abilities and skills. These abilities are not mutually exclusive and a combination of several abilities and/or skills might be particularly promising. However, performing poorly on one dimension may not be counterbalanced by scintillating on another dimension. Insufficient academic performance for instance will not be

compensated for by additional extracurricular activities or vice versa. Accordingly, different ability signals will not represent substitutes, but rather complements:

Hypothesis 4: Different ability signals will not have a substitutive, but a complementary effect on awarding probabilities. Therefore, reinforcing interaction effects of different ability signals will occur.

Assuming a perfectly rational decision process made by scholarship evaluators, non-productivity-related applicant characteristics such as gender or race are not likely to have an impact on selection decisions. However, the hiring discrimination literature has shown that recruiters do not decide perfectly rationally and discriminate against certain applicants due to several reasons. Discrimination against certain minority groups occurs as a consequence of either taste-based or statistical discrimination as has been demonstrated in Chapter 4.2.3.1. Albeit, in a stipend awarding process several of the theoretically anticipated sources of (hiring) discrimination are simply non-existent. Out of the three distinct sources of taste-based discrimination (employer, employee and customer discrimination), only one is applicable in the current context. As neither 'co-workers' nor 'customers' exist in a scholarship context, the only rationale behind taste-based discrimination could be the evaluator's own distaste against certain minority candidates. Statistical discrimination in turn might occur whenever an evaluator has previously made an unpleasant experience with other students belonging to the same minority group and – to the best of their knowledge and in the absence of further information – infers a lower generalized ability for all affiliates of this minority group. Examples of unpleasant experiences of this kind can be either general (e.g. performing poorly in university or never being punctual) or stipend specific (e.g. dropping out of the program although being awarded the stipend). But as evaluators usually do neither meet successful applicants nor get any information on how actual awardees perform after having been selected, they should not observe group differences in performance. In addition, as a lot of other ability-related information is provided in the application, e.g. grades, certificates etc., uncertainty about applicant quality should be rather low. Consequently, in the very unlikely case of discrimination, less favorable treatment treatment occurs due to evaluator's idiosyncrasies, but should not occur in the aggregate of all selection decisions or on average. However, in order to be able to empirically test the occurance of discrimination in the present study, Hypothesis 5 states as follows:

Hypothesis 5: Applicant characteristics (indices) will affect the selection probability, i.e. discrimination will occur.

Some types of indices however are expected to have an influence on selection decisions without being indicative of discrimination in selection. Age, personality and social background of an applicant for instance are expected to affect evaluator decisions indirectly. Younger students will (c.p.) be preferred to older students as being young and having achieved the same as an older student will most likely be interpreted as an indirect ability signal reflecting higher motivation. Similarly, in concordance with elite theory, individuals from higher social classes will c.p. have higher chances to be granted the scholarship as they do not only possess increased human, but also social capital which can be especially advantageous in final selection stages (interview). Nevertheless, this again cannot simply be interpreted as discrimination against working class applicants. To a greater degree, originating from a high socio-economic background leads to increased rhetoric skills which in turn represents a signal for increased aptitude.

Analogous to applicant characteristics, rater characteristics are not expected to have any effect on rational selection decisions. In personnel selection, it has indeed been shown that female evaluators are more lenient in evaluating candidates on several work-related outcomes, but in terms of final selection decisions this leniency usually 'disappears'. Furthermore, increased interview structure has been demonstrated to rule out the influence of individual rater characteristics on selection probabilities. In order to be able to falsify the assumption of any rater characteristic effect, Hypothesis 6 is developed:

Hypothesis 6: Selection probabilities will be dependent of individual rater characteristics.

In the same vein, the influence of further social and situational factors such as applicant-rater similarity, panel composition and interview time is expected to be non-significant, but needs to be tested in order to reduce a possible omitted variable bias.

Hypothesis 7: Applicant-rater similarity will lead to more favorable ratings.

Hypothesis 8: Situational or extraneous factors will affect selection decisions.

In general, pre-selection decisions are expected to be more predictable than final selection decisions (based on written applicant credentials) as the amount of unobservable factors and hence the noise included in the model increases in final selection decisions. Hypothesis 9 accordingly states

Hypothesis 9: The amount of explained variance in pre-selection decisions will be greater than in final selection decisions.

PART C: EMPIRICAL EVIDENCE ON THE DETERMINANTS OF SUCCESSFUL SCHOLARSHIP APPLICATIONS

6 THE DATA SET

6.1 METHOD AND DATA

In order to address the previously established hypotheses and research questions, an empirical analysis of actual stipend applications and awarding decisions was conducted. One single, sample scholarship program was investigated and it was observed and analyzed in retrospect, who was actually a) invited to an interview after the sifting process and b) finally awarded the scholarship.

As has been discussed in the theoretical part, analyzing real decision processes yields superior results to both experiments and surveys (in terms of external validity). Only by examining real decisions made in the field, the researcher is able to reveal true preferences that evaluators or recruiters are not likely to admit in e.g. surveys (revealed versus stated preferences). In field experiments, however, only a limited number of variables of interest can be actively manipulated and tested. As previous empirical findings on stipend awarding decisions are rare, it was not possible to select appropriate experimental treatment variables. Instead, it is the purpose of this thesis to provide in-depth insight into various previously unknown signaling and screening activities in scholarship selection processes.

The data set used for answering the research questions was provided by the 'Deutsche Akademische Austauschdienst' (DAAD), an independent German exchange service which regularly awards scholarships to academics of any degree. In concordance with functional elite theory, the organization's mission is to award stipends solely on the basis of individual performance:

> *„Even in controversial times, the DAAD is committed to an elite sponsorship which is purely performance-oriented, secured by independent academic committees and hence accessible to everybody who complies with these high standards. Apart from intellectual abilities, stipend awardees need to possess a personality profile which gives reason to expect that the stipend awardee will directly or indirectly return the favor and pay the sponsoring society something back from what he or she has earned with the aid of the stipend."*[83]

[83] Own translation from http://www.daad.de/portrait/wer-wir-sind/programme/08941.de.html.

It can be easily revealed from this mission statement that the present stipend program represents a purely merit-based scholarship program. Stipends are not awarded as a function of a student's financial need or socio-economic status (as it would be the case in means-tested scholarship programs), but based on individual achievement only.

The applications all stem from the same subdivision of this exchange service and two subsequent application periods[84] have been examined. In the investigated sample, all applicants are either German (i.e. possessing the German citizenship) or educational residents in Germany (meaning they have obtained their high school diploma in Germany or at a German high school). Non-residents are not able to apply for this specific kind of scholarship and are therefore not represented in the sample. All of the applicants for this particular scholarship program (called 'Germans to North America') are undergraduate students who wish to spend two semesters (i.e. approximately nine to ten months) at a North American university and hope to be financially supported by the institution.[85]

The application process

In order for an application to be considered, students need to hand in at least[86] the following documents

- Application form (including a photograph),
- Typed complete curriculum vitae (CV), including course of studies,
- Detailed description of the curriculum to be studied abroad (max. 5 pages),
- Reference letter and standardized evaluation form of a faculty member,
- Table of previous academic achievements at university,
  (Copies of grade certificates, intermediate examination certificates or diplomas),
- Copy of high school diploma (including grades for individual subjects) as well as a
- Language certificate (either TOEFL or another approved certificate).

[84] The investigated scholarship periods, i.e. the period in which applicants planned to study abroad, were the academic years 2008/09 and 2009/10. Stipend awarding was decided upon one year prior to departure , i.e. in 2007 for 2008/09 and in 2008 for 2009/10.

[85] The financial assistance granted to those awarded the scholarship is substantial and should cover most of the expenses the students have during their stay abroad. It adds up to 850€ per month plus a fixed monthly medical insurance rate plus tuition fees up to 15,000€/year per person.

[86] Further certificates or information handed in by the applicant were also forwarded to the evaluators.

All complete written applications handed in before the expiration of the application deadline are distributed among several individual pre-selection evaluators who are asked to invite approximately 50% of all applicants to an interview. Pre-selection evaluators are professors teaching in diverse fields at different German institutions who volunteered to be part of this selection process. Applications are distributed among different evaluators on the basis of the applicant's distinct field of studies which should be identical or at least related to the evaluator's field of research. Evaluators are not bound to assess specific KSAs of applicants, but are simply asked to invite the most promising of all candidates. As stated in the organization's mission, stipend awardees should possess appropriate intellectual abilities as well as a promising personality profile. In order to justify their decision and to be able to compare different applicants, evaluators are asked to assign each candidate a pre-selection score ranging from 0-100. Actually, this pre-selection score is subdivided into three differently weighted categories in order to provide evaluators with some guidelines: Pre-selection evaluators are asked to assign 0-55 (out of the 100) points for academic qualification, 0-15 points for extracurricular qualification and 0-30 points for the specific project (goals and preparation of the stay abroad). Usually, all applicants receiving a score ranging from 80 to 100 are subsequently invited to an interview. Therefore, in addition to the documents handed in by the applicant, the data set also comprises distinct evaluation forms stating each candidate's individual and total pre-selection scores and whether or not the applicant was invited to an interview and therefore further considered in the application process. Additionally, information about the pre-selection evaluator (e.g. gender, position, age, and field) is available.

Whenever applicants have successfully 'survived' the sifting phase, they have been invited to a personal interview in front of a committee. In this case, further information about the interview situation and structure (day, time & length), the evaluation committee (size, gender and age composition) and the outcome of the selection interview (final score and scholarship awarded: yes or no) has been recorded and can be used for analysis.

All of the variables that could be extracted from both paper applications and pre-selection or interview notes and used for further analysis will be presented in detail in the following subsection (Descriptive Statistics).

6.2 DESCRIPTIVE STATISTICS

Initially, the data set contained 504 complete scholarship applications.[87] These represent all of the applications handed in for this specific scholarship program in the two application periods examined. 243 of all applications were handed in for the first selection round conducted in 2007 (scholarship period: 2008/09) and the remaining 261 applications were decided upon in 2008 (scholarship period: 2009/10). Complete paper files were made accessible to the researcher so that numerous (possible) independent variables could be extracted from the above mentioned paper documents.

6.2.1 The Applicant Pool

6.2.1.1 Applicant Characteristics

48.6% of all applicants (245) were female and 51.4% (259) were male. In terms of the application decision made by the individual student, no gender differences could hence be observed and both female and male students applied for a study-abroad scholarship in (almost) equal shares. At the time of application, i.e. one year prior to starting their semester abroad, applicants were on average 21.9 years old. The youngest of all applicants was 19, the oldest 32. In terms of duration of study, the average applicant had already studied for 4.4 semesters when applying for the stipend, but again the range between the shortest (2nd semester) and longest (14th semester) length of study was substantial. Most of the applicants (79.4%) desire to study in the United States of America while 20.6% plan to spend their year abroad in Canada.[88] 65.5% of all applicants indicate that they will need to pay tuition at their guest institution.[89] Whenever tuition fees have to be paid by the student, they on average add up to US$ 21,953[90] for the entire stay abroad, but again tuition fees are not uniformly distributed. Some applicants only expect to pay US$ 1,570 whereas others envisage tuition fees as high as US$ 80,000. 139 applicants (27.8%) indicate to take part in an organized study-abroad program offered by their home institution. Most of the applicants (97%) plan to attend lectures in English language. The remaining 3% of all applicants intends to study in French language (at Canadian universities only).

---

[87] However, not all of the applications could be used in the subsequent multivariate analysis due to missing values in important explanatory variables (see Chapter 7 for details).

[88] Note that all of the applicants in this specific program want to study abroad in North America. Students wanting to get financial assistance for a stay elsewhere would apply for another program in another subdivision of this organization.

[89] Although all North American institutions require tuition fees, tuition can be waived in some cases. Examples for these exemptions are bilateral agreements between partner universities or entire federal states (e.g. Baden-Wuerttemberg and Ontario).

[90] Tuition fees indicated in € or CAD have been transformed into US$ for the ease of interpretation and comparison.

In terms of previous sponsorships, 22.6% of all applicants indicate to receive BAföG.[91] 74 candidates, i.e. 14.7% of all applicants, indicate to have been awarded another merit-based stipend before. Most of these previously awarded candidates (30) indicate to be a scholarship recipient of the 'Studienstiftung des Deutschen Volkes', the most renowned of the eleven 'Begabtenförderungswerke' in Germany. Other organizations that have previously supported candidates include 'e-fellows.net' (9 applicants) and the 'Konrad-Adenauer-Stiftung' (7 applicants), but also specific programs offered by the respective home institution have been mentioned (7 applicants). 212 of all 504 candidates (42.1%) additionally state in their application that they have also applied for other scholarships offered by other institutions promising to support their planned stay abroad.

The applicants are students of 64 different higher education institutions in Germany which are almost all public institutions (98%). More than three quarters of all applicants are enrolled at a German university (76.4%). 15.4% of all candidates indicate to study at a technical university and the remaining 8.2% of all applicants are enrolled at a university of applied sciences.[92] The home institution of 42.8% of all applicants is located in the German state of Baden-Wuerttemberg and 13.6% indicate to be currently enrolled at a Bavarian tertiary education institution. Only 10% of all applicants study at an institution located in North-Rhine-Westphalia, the most heavily populated of all German states. Very few applicants (5.2%) are enrolled at an institution which is located in one of the new Eastern states of Germany. 37.3% of all applicants study a subject counted among 'Law, Economics or Social Sciences' and 28.4% are categorized as students of 'Linguistic and Cultural Sciences'. The disciplines 'Engineering' and 'Mathematics, Informatics and Natural Sciences' are represented in the sample 71 times (14.1% of all applications) and 85 times (16.9%), respectively. The remaining 17 applicants are enrolled in 'Arts', 'Medicine' or interdisciplinary fields.[93]

In terms of educational attainment, especially high school and university grade averages need to be mentioned. The average scholarship applicant has passed secondary school with a grade average of 1.8 (range: 1.0 – 3.7) and has achieved university grades averaging 1.9 (range: 1.0 – 3.7). It needs to be mentioned here that the German grade system is different from the American grading system for instance. In Germany, 1.0

[91] A German nationwide means-tested sponsorship especially developed for students whose parents could otherwise not afford sending their children to university or school. For more details on this needs-based program, see Chapter 2.

[92] University of applied sciences = '*Fachhochschule'.*

[93] Due to the peculiarities of the study of art, the German Academic Exchange Service offers separate programs especially designed for art students. This fact might explain the relatively low share of art students in the present sample.

represents the best grade (reflecting excellent performance) and 6.0 the worst grade (reflecting insufficient achievement). As it is reverse coded, a 1.0 therefore corresponds to an A in the American system, a 2.0 to a B, a 3.0 to a C, and so on. In order to pass a course, a student needs to achieve at least a 4.0 which means the grade needs to be lower than or equal to 4. This grading system is utilized both in high school and university. Accordingly, averages of 1.8 and 1.9 respectively correspond to a B(+) in the American system.

In their CVs, 84.3% of all applicants indicate some kind of extracurricular activity (ECA). These activities can be classified into several categories which are spread among all applicants as follows[94]: 48.2% of all applicants are involved in the organization of youth, sport or recreation activities, 24.8% of all applicants voluntarily work for a social services provider whereas 22.8% put effort into cultural activities. 17.7% indicate to pursue a political and 14.3% a clerical ECA. 21.2% engage in ECAs related to their student association or faculty and 14.5% have been a member of the student representation in high school. Finally, 8.1% voluntarily support exchange students at their home institution.

As requested by the scholarship granting organization, applicants provide both free-form recommendation letters and a standardized evaluation sheet filled in by the recommender, usually a faculty member of the home institution. Only one recommendation letter and evaluation sheet was required, but several applicants handed in multiple LORs. On average, 1.3 LORs were handed in which were all very generous: On a scale from 0-10 (0 reflecting not at all and 10 perfectly suitable), the average applicant achieved 9.25 points. In addition, 83.4% of all recommenders indicated to know the recommendee well.

In Germany, written applications frequently contain additional information which are not requested by e.g. employers or evaluators, but are provided commonly on a voluntary basis. Statements about the applicant's parents count among this additional information. In the current applicant pool, 33.5% of all applicants voluntarily mentioned their parents, or more precisely, their parents' occupation and status. Of the 169 applicants who provided information about their parents, 70.4% come from an academic parental home, i.e. based on the indicated current occupation or academic degree, at least one parent has presumably graduated from university.

[94] As numerous ECAs can be pursued by any individual applicant, multiple answers were possible and the percentages do not add up to 100.

6.2.1.2 Comparison to the Entire German Student Body

In order to correctly interpret the previously presented characteristics of all applicants, the applicant pool needs to be compared to some kind of reference category. In the present case, the average German student represents an appropriate reference.

As has been discussed in the model development section, the individual decision to apply e.g. for a scholarship precedes the actual application. Only when a student's, i.e. a potential applicant's, utility of applying is greater than the utility of not applying, he or she will make the effort of collecting all the necessary documents and handing them in. Hence, it is anticipated that those students who actually decided to apply for the scholarship (and whose application we are consequently able to analyze) differ significantly from the average student. If this is indeed the case, the applicant pool represents a (positively) self-selected group of students which will presumably not only be more self-confident, but also more qualified (e.g. in terms of grades) than the average German student.

The following table contrasts some of the characteristics that could be observed among all applicants with the characteristics of an average German student. With the help of this comparison, we can find out whether the applicant pool represents a distinct, (positively) self-selected group of students or whether applicants do not differ significantly from 'usual' students in Germany. Information about the average student's characteristics was obtained from several official sources including Isserstedt et al. (2007), Middendorff, Isserstedt and Kandulla (2009) as well as various publications from the German Federal Bureau of Statistics (Destatis 2009, 2011c; Schmidt 2009). Of course, this comparison is only descriptive in nature and does not control for confounding effects, but nevertheless sheds some light on the scholarship application decision of German students.

Variable of Interest	Average Value of …	
	Applicant Pool	German Student Body
Study-Abroad Intentions/Experience		
Percentage of students with study-abroad intentions/exper.	100%	15%[95]
Academic Achievement		
High School Grade Average	1.80	2.2[96]
University Grade Average	1.94	2.2[97]
Type of Home Institution		*Data from Destatis (2009)*
University	76.4%	66.3%
Technical University	15.4%	*n.a. (incl. in Universities)*
University of Applied Sciences	8.2%	29.6%
Location of Home Institution		*Data from Destatis (2009)*
Baden-Wuerttemberg	42.8%	12.9%
Bavaria	13.6%	12.8%
Berlin	7.4%	6.7%
Brandenburg	1.0%	2.3%
Bremen	1.2%	1.6%
Hamburg	0.4%	3.6%
Hesse	3.0%	8.5%
Lower Saxony	5.8%	6.9%
Mecklenburg-Hither Pomerania	0.2%	1.8%
North Rhine-Westphalia	10.0%	23.9%
Rhineland-Palatinate	7.6%	5.3%
Saarland	2.0%	1.1%
Saxony	3.2%	5.3%
Saxony-Anhalt	0.8%	2.6%
Schleswig-Holstein	0.8%	2.4%
Thuringia	0.0%	2.5%
Other	0.2%	-
Home Institution in New Eastern State (Berlin excl.)	5.2%	14.5%
Work Experience		*Data from Isserstedt et al. (2007)*
Student completed vocational training prior to studying	5.6%	25%
Previous Sponsorships		
BAföG	22.6%	25.5%[98]
Any Merit-Based scholarship	14.7%	1.1%[99]
'Studienstiftung des Dt. Volkes'	6.0%	0.5%[100]
Extracurricular Activities		*Data from Fischer (2006)*
Share of Students Pursuing an ECA	84.3%	66.7%
Applicant Characteristics		*Data from Destatis (2009)*
Share of Female Students	48.6%	47.8%
Share of Students Born in Germany	91.8%	*n.a.*
Average Student Age	21.9 years	25.3 years

Table 6-1 Comparison of Applicant and Average German Student Characteristics

[95] Isserstedt et al. (2007), 166.
[96] Data obtained from the 'HIS-Studienberechtigtenbefragung' 2006-2010 (HIS 2012).
[97] Wissenschaftsrat (2007), 32.
[98] Schmidt (2009), 168. Calculation: 494,480/1,941,763 students in 2007 (Destatis 2008).
[99] Middendorff, Isserstedt & Kandulla (2009), 14 (Calculation: 20,794/1,941,763 students in 07).
[100] Middendorff, Isserstedt & Kandulla (2009), 14 (Calculation: 8,717/1,941,763 students in 07).

First and foremost, all students having applied for the examined scholarship program definitely intend to study abroad for several months. This fact alone reflects a certain self-selection: Only 15% of all advanced students[101] in Germany report to have ever studied abroad for some time (Isserstedt et al. 2007). Although – to the best of my knowledge – no comprehensive study exists which has empirically tested the determinants of a study-abroad decision, it might be assumed that those actually studying abroad differ significantly from the average student. With the help of surveys, some researchers have tried to shed some light on the determinants of studying abroad. Heublein et al. (2011) for instance have directly asked a sample of advanced students for their reasons to go or not to go abroad during their studies. Acquiring new experience, getting to know other cultures, improving their language skills and boosting their career opportunities count among the most frequently indicated motives for studying abroad. Frequently mentioned arguments against a semester abroad are financial difficulties, long separation from family and friends, the organizational effort needed to prepare such a stay, loss of time, low compatibility of study programs and difficulties in obtaining information about study-abroad possibilities (Heublein et al. 2011). From these answers it can be inferred that internationally mobile students indeed differ (positively) from the average student, e.g. in terms of openness to experience, determination, motivation, assertiveness and simply organizational skills. In terms of academic achievement, some differences between the applicant pool and the German student population are apparent. Whereas all stipend applicants exhibit high school grade averages of 1.8 and university grade averages totaling 1.94, the average German student has 'only' achieved a 2.2 (HIS 2012) and a 2.2 (Wissenschaftsrat 2007) respectively. Hence, especially students with particularly good grades seem to apply for a scholarship. This fact might on the one hand reflect the applicants' anticipation of certain selection criteria: As students expect grades to play an important role in the selection process, only those with 'adequate' grades apply. Similarly to what Manski and Wise (1983) found for college applications, applicants try to anticipate the selection criteria used by the selecting organization. On the other hand, the explanation for the difference in academic achievement could be that all students planning to go abroad differ from the average student also in terms of grades. It is feasible that only students with better grades – as an indicator for high levels of determination, motivation and diligence – decide to go abroad and consequently, the applicant pool rather resembles the group of all students going abroad.[102]

[101] Advanced student: In the 6th (university of applied sciences) or 8th (university) semester.

[102] However, no data on the grade distribution of internationally mobile and immobile German students exists that would allow for a more detailed investigation.

Although almost one third of all German students are enrolled at a university of applied sciences, only 8.2% of all applicants study at this type of institution. This proportion most closely corresponds to the share of students from universities of applied sciences among all scholarship recipients in Germany: In their survey, Middendorff, Isserstedt and Kandulla (2009) have found that only 8% of all actual scholarship recipients are enrolled at a university of applied sciences. In the present case, two distinct explanations may exist for the observed imbalance: Either students from universities of applied sciences decide to go abroad to a lesser extent than students from a 'regular' university or students enrolled at these institutions do go abroad, but do not apply for a scholarship. Reasons for their reluctant application behavior could be a lower self-esteem in comparison to university students or a mere paucity of information about these stipend possibilities. Looking at the results of Middendorff, Isserstedt and Kandulla (2009) who have examined unconditional scholarships, i.e. not attached to a certain program or project, one can find support for the reluctant-application-behavior hypothesis. On the other hand, as the Isserstedt et al. (2007) survey data show, the share of students from universities of applied sciences possessing study-abroad experience is indeed substantially lower than for university students: Whereas 19% of all university students have made some international experience during their studies, only 8% of all university-of-applied-sciences students have made this experience. Possible reasons for their study-abroad reluctance are on the one hand their predetermined curriculum not allowing for any delays and on the other hand their (on average) lower socio-economic background leading to increased financial constraints (Isserstedt et al. 2007). However, this again would be an argument for an increased share of applicants for a stipend promising to financially support study-abroad projects. But on the other hand, students from universities of applied sciences might have other means of financial support at their disposal: means-tested scholarships such as BAföG which are only available for students from less affluent families or corporate stipends from companies collaborating closely with these kinds of institutions. Hence, various explanations for the type-of-institution differences are plausible, but it needs to be emphasized that only a minority of all applicants study at a university of applied sciences.

One of the most striking differences between the applicant pool and the entire German student body is the distribution of home institution locations across Germany. This differential distribution is most apparent for the federal states of Baden-Wuerttemberg and North Rhine-Westphalia, as can be derived from figure 6-1.[103] Whereas only 12.9% of all German students study at an institution located in the federal state of Baden-Wuerttemberg (Destatis 2009), more than two fifths (42.8%) of all applicants are enrolled at an institution located in this state. On the contrary, only 10% of all applicants come from an institution located in North Rhine-Westphalia although almost one quarter of all German students (23.9%) study at an institution located in this federal state (Destatis 2009). Additionally, only very few applicants (5.2%) are from one of the German institutions located in the new Eastern states (Berlin excluded) although 14.5% of all German students are enrolled at these institutions (Destatis 2009).

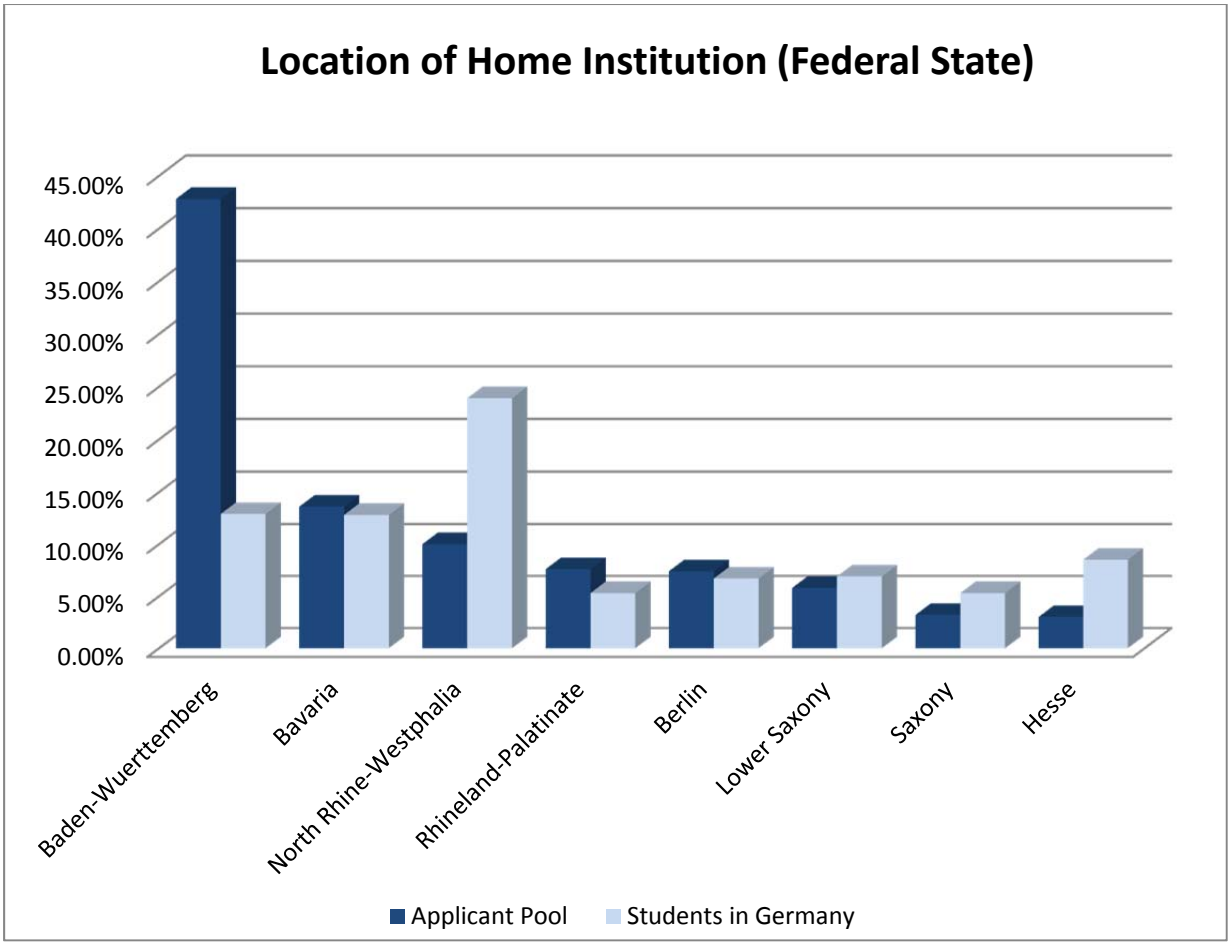

Figure 6-1 Location of Home Institution: Applicant Pool versus Student Population
Sources: Destatis (2009) and Own Data Set

[103] Note that for the ease of illustration only selected federal states are presented here.

Several explanations for this substantial imbalance are conceivable. First of all, students from Baden-Wuerttemberg could be extraordinarily eager to study abroad for some time. However, no empirical data to test this hypothesis is available. Heublein et al. (2011) for instance did not differentiate between students from varying federal states. Isserstedt et al. (2007) did not indicate state-specific study-abroad rates either.[104] Another possible reason for the increased share of Baden-Wuerttemberg students in the applicant pool would be the superior academic 'quality' and hence an increased self-confidence (and application success expectation) of these students. Indicators for academic quality could either be the individual grade average or the ranking results of the universities they are attending. The comprehensive overview of university grades published by the Wissenschaftsrat (2007) however does not differentiate between students from different federal states. In terms of high school grades however, such a differentiation is made by the KMK (2006): The average Baden-Wuerttemberg high school student indeed achieves better grades than those in most other federal states – only high-school graduates from Thuringia are slightly better. However, this difference does not seem to be substantial in comparison to high school graduates in Lower Saxony and Berlin who have achieved the worst grade averages. In the applicant pool however, students from Baden-Wuerttemberg indeed demonstrate significantly better grades than the remaining applicants: In comparison to students from all other federal states, they have achieved high school grade averages of 1.6 (other states 1.9; $p < .001$) and university grade averages of 1.9 (other states 2.0; $p < .05$). Similarly, in terms of home institution quality – measured by both CHE-research-reputation and study-situation rankings[105] – Baden-Wuerttemberg applicants also study at higher quality institutions (see table 6-2)[106].

Home Institution Ranking	Federal State of Home Institution= Baden-Wuerttemberg?		Difference	Significance
	Yes	No		
Study Situation	2.18	2.58	-.40	***
Research Reputation	0.29	0.16	0.13	***

Table 6-2 Home Institution Reputation: Baden-Wuerttemberg Applicants versus Other

[104] Only the study-abroad behavior of students having graduated from high school in 'new' and 'old' federal states of Germany was compared, but no significant differences were found.

[105] For an explanation of how the CHE-ranking is calculated see table 6-7 on p. 121 or visit www.che-ranking.de.

[106] Unless otherwise specified, all subsequent tables and figures are own illustrations based on the data set collected during this research project.

This fact might indirectly, i.e. via an increased self-confidence, encourage them more to apply for a merit-based stipend. Furthermore, students from these institutions are likely to represent a positively self-selected group as only the best students decide to study at high-reputation institutions in the first place.

In addition to quality advantages, students from Baden-Wuerttemberg might simply be more acquainted with the specific stipend possibility offered by the investigated organization. As we are not able to directly test the degree of familiarity with the particular program, a proxy for program prevalence is needed. The number of evaluators teaching at institutions located in each federal state might serve as a proxy, as it is feasible that evaluators are actively promoting the program at their home institutions. However, only two out of 30 evaluators have been teaching at a Baden-Wuerttemberg university which might not explain the increased fraction of Baden-Wuerttemberg applicants. On the contrary, six evaluators came from North Rhine-Westphalia, the federal state where comparatively few applicants were enrolled. If the degree of familiarity with the stipend possibility had an influence, North-Rhine Westphalian students were even expected to apply more frequently than they did, as the scholarship granting organization investigated here is headquartered in Bonn, North Rhine-Westphalia. However, only 10% of all applicants study in this federal state of Germany. Reversing the chain of reasoning presented above for the case of Baden-Wuerttemberg, we would expect lower quality students and institutions to be located in North Rhine-Westphalia. As no comprehensive overview on university quality in different federal states exists, we again use applicant and their university's quality as a proxy for overall academic quality. Applicants studying at a North Rhine-Westphalian institutions indeed have only achieved high school grade averages of 2.0 which are significantly worse than those of all other students (1.77, $p < .05$), but in terms of university grades they do not differ from the rest of the applicant pool. Home institution quality measured as an institution's research reputation is also significantly lower for North Rhine-Westphalian applicants (see table 6-3).

Home Institution Ranking	Federal State of Home Institution= North Rhine-Westphalia?		Difference	Significance
	Yes	No		
Study Situation	2.51	2.41	.10	+
Research Reputation	0.15	0.22	-.07	**

Table 6-3 Home Institution Reputation: North Rhine-Westphalian Applicants versus Other

Looking at table 6-1, another location-related difference becomes obvious: Only comparably few students from institutions located in one of the Eastern federal states of Germany apply for a scholarship. Again, one possible reason might be a lower study-abroad rate among students from these states. However, Isserstedt et al. (2007) showed that the study-abroad rate does not vary for students from 'new' and 'old' federal states. As the average income per capita in the new states is lower than in the rest of the Federal Republic of Germany (VGRdL 2011), one would expect students in these regions to be more in need of scholarships. However, they could be more reluctant to apply for a merit-based stipend as they are more frequently entitled for a (non-competitive) means-tested scholarship. Indeed, 19.6% of all BAföG recipients study in one of the new states (Destatis 2011c) although 'only' 14.7% of all students are enrolled at institutions located in Eastern Germany. However, this difference does not fully account for the observed differential application rates for 'Eastern' and 'Western' students. Other factors such as institution and applicant quality[107], student self-confidence, institutional support and stipend tradition might also play a role in determining the individual decision to apply for a merit-based scholarship.

Another salient difference can be observed in the share of students who have completed vocational training prior to their studies. Whereas one quarter of all German students has already completed an apprenticeship before studying, only 5.6% of all applicants possess this kind of work experience. Again, several explanations for their reluctance to apply are feasible. It is worth testing whether students having completed an apprenticeship prior to studying go abroad less frequently than students who have started studying directly after completion of high school. Reasons for this could be increased financial constraints and the expected time loss which might both be more important to students who have already completed three years of vocational training. This in turn leads to a lower individual marginal utility of studying abroad. However, no data on this particular decision are available – Heublein et al. (2011) for instance did not differentiate between students with or without completed apprenticeships. Secondly, one might argue that a delayed decision to attend university, i.e. only after having completed a 3-year apprenticeship, can be interpreted as a sign of insufficient determination. These students could fear that their past decisions might be interpreted as hesitancy which in turn leads them to refrain from applying for scholarship programs. Due to missing data to test the assumptions, however, these arguments remain speculative in nature.

[107] A comparison of applicant grades as well as institution quality (ranking) was not possible due to the small number of observations for institutions located in Eastern federal states.

Finally, applicants also differ from the average student in terms of previous merit-based scholarships, the amount of extracurricular activities and age. A substantially higher fraction of previously awarded students exists among applicants (14.7% versus 1.1% of all students). Presumably, having been awarded another stipend before increases a student's self-esteem and encourages him or her to also apply for other programs. In terms of ECAs, 84.3% of all applicants indicate to pursue at least one of these activities while 'only' two thirds of all German students do so (Fischer 2006). Again, applicants might assume e.g. from the stipend-granting organization's mission that ECA will constitute an important selection criterion. Whether or not this expectation causes (potential) applicants to start pursuing (or simply mentioning) an ECA or whether only students who are active anyway apply for a scholarship, cannot be answered in this context.

In terms of applicant characteristics, no differences in gender or origin between the two groups (all students vs. all applicants) could be observed. However, it is salient that applicants are 3.4 years younger than the average student. This could again be an indicator for increased determination and motivation, but it could also be caused by the simple fact that applicants for a study-abroad scholarship are usually still at the beginning of their studies and have not completed apprenticeships prior to studying.

Summarizing, the applicant pool does not represent a random sample of average German students. Only particular students seem to decide to apply for a scholarship and several self-selection effects occur prior to applying for a scholarschip program.

6.2.2 Pre-Selection Situation

In the first selection round, 30 different evaluators were in charge of assessing the applicants' potential and were asked to invite approximately 50% of all applicants to an interview. As mentioned earlier, all evaluators were professors from several German institutions. The average evaluator age was 54 years: The youngest evaluator was aged 33, the oldest 68 at the time of selection. Only six of all evaluators, i.e. 20%, were female. This almost exactely corresponds to the share of female professors in Germany which amounted 19.2% in 2010 (Destatis 2012b). On average, one evaluator had to decide upon 24 applications during this first selection round. Again, the range was substantial: Some evaluators only had to evaluate two applications whereas others needed to choose between 40 different applicants.

In terms of pre-selection outcomes, two measures are available for each candidate. The pre-selection score on a scale from 0-100 and the binary decision whether or not the applicant is invited to an interview. The average pre-selection score totals 75.4 and ranges from 19.5 to 100. In total, 54% of all applicants, i.e. slightly more applicants than the 50% desired by the organization, passed the first selection round and were subsequently invited to an interview. This percentage was consistent for both examined periods: In 2007, 131 out of 243 applicants (53.9%) and in 2008, 141 out of 261 applicants (54%) 'survived' the first selection round. One peculiarity in terms of evaluator gender however is salient. When we observe pre-selection success quotas separately for male and female evaluators, a certain female leniency in pre-selection decisions becomes obvious.

Invitation	Evaluator Gender		Total
	Male	Female	
Yes	215 (52,57%)	57 (60,00%)	272 (53,97%)
No	194 (47,43%)	38 (40,00%)	232 (46,03%)
Total	409 (100%)	95 (100%)	504 (100%)

Table 6-4 Invitation Quotas according to Pre-Selection Evaluator Gender

As can be derived from table 6-4, female evaluators even suggest inviting 60% of all applicants to an interview although being asked by the organization to sort out half of all applicants in this first selection round. Male evaluators also slightly exceed the limit as they on average invite 53% of all applicants they are asked to assess.

6.2.3 Final Selection Situation

Only 254 of the 272 applicants which were invited to a selection interview actually attended the interview.[108] Individual selection interviews took place from 9 am to 7 pm on three consecutive days in four parallel committees and interview times were distributed as presented in table 6-5.

Interview beginning between	Frequency	Percentage
09 am to 09:59 am	34	13.4 %
10 am to 10:59 am	38	15.0 %
11 am to 11:59 am	40	15.7 %
12 pm to 1:30 pm	39	15.4 %
LUNCH BREAK	----------*LUNCH BREAK*----------	
2 pm to 2:59 pm	37	14.6 %
3 pm to 3:59 pm	22	8.7 %
4 pm to 4:59 pm	21	8.3 %
5 pm to 6:59 pm	23	9.1 %
Total	**254**	**100.00 %**

Table 6-5 Interview Times

Applicants had to present themselves in front of a selection committee (panel interview) composed of three to seven evaluators (average number of panel members: 4.35). Committees were mainly built on a subject-specific basis which means that evaluators and applicants in general should teach and study in related areas. Most of these evaluators had also been involved in the preceding pre-selection assessments and only very few evaluators were appointed for interview selection only. The average share of female evaluators in the panel amounted to 26%, but some committees were also completely composed of female or male evaluators. Therefore, the share of female evaluators ranged from 0 to 100%. In 82% of all cases however, applicants were confronted with an interview panel predominately consisting of male evaluators, i.e. the majority of evaluators were men. On average, evaluators were aged 52.4 years (range:

[108] Most of the 18 students who did not show up declined the offer on the basis that they had either dropped their study-abroad-plans or had already been awarded another scholarship. It needs to be mentioned that these 18 students differ slightly from the remaining 254 applicants who have actually attended the interview. For instance, the 'no shows' more often attend a technical university and have more frequently managed to be supported by another merit-based stipend program (esp. 'Studienstiftung'). Furthermore, each of these 18 students pursues at least one extracurricular activity. On the other hand, relatively few of these 'no shows' have handed in a TOEFL. In terms of their study abroad plans, comparably many of the 'no shows' planned to attend a THE Top10 guest institution. However, only very few of them wanted to attend an institution in the US-Northeast, but rather planned to study in the US-West. Comparatively many interview annulations occurred in the stipend period 2009/10. A more detailed comparison of descriptive statistics for the two groups ('no shows' vs. 'interviewed applicants') is provided in Appendix 1.

43.5 to 62 years). This again closely corresponds to the average age of all university professors in Germany which amounted 51.2 years in 2011 (Destatis 2012c).

The average selection interview lasted 14 minutes, but some applicants were only interviewed for three minutes whereas others presented themselves for more than 20 minutes. After having interviewed an applicant, the committee discussed his or her aptitude and agreed upon a final selection score on a scale from 0 to 100. Again, scores above 79 led to a positive decision, scores below that threshold resulted in a rejection of the candidate. Final scores averaged 80.2 points and ranged from 60 to 95. In total, 57.5% of all interviewees were awarded the stipend in the end. As presented earlier, the percentage of applicants invited to an interview did not differ significantly between the two selection periods, but the ratio of subsequently awarded scholarships did: Whereas 64% of all interviewed applicants were awarded a grant in 2008/09, only 51% of the interviewees in 2009/10 managed to receive a scholarship.

Stipend awarded?	Stipend Period		Total
	2008/09	2009/10	
Yes	80	66	146
	(64,00%)	(51,16%)	(57,48%)
No	45	63	108
	(36,00%)	(48,84%)	(42,52%)
Total	125	129	254
	(100%)	(100%)	(100%)

Table 6-6 Stipend Awardings according to Stipend Period

The main reasons for this difference are variations in budget constraints between the two distinct selection periods. Hence, during the multivariate analysis of selection success, the selection year always needs to be included as a control variable.

6.2.4 Overview of Descriptive Statistics and Operationalization of Variables

In order to give a comprehensive overview of all the variables extracted from applications and selection contexts, table 6-7 summarizes all variables of interest for the subsequent multivariate analysis of the scholarship selection process. This list does not only include applicant and evaluator characteristics, but also contextual factors.

Variable	Operationalization	# of Obs.	Mean	SD	Min	Max
DEPENDENT VARIABLES						
Pre-Selection Score	On a scale from 0 to 100	504	75.42	14.08	19.5	100
Invitation to Interview	Dummy (Yes=1, 0 otherwise)	504	.540	-	0	1
Final Selection Score	On a scale from 0 to 100	254	80.23	6.07	60	95
Scholarship Awarded	Dummy (Yes=1, 0 otherwise)	504	.290	-	0	1
INDEPENDENT VARIABLES						
Academic Achievement						
High School Grade Average	Grade average in German high school system (1.0 being the best & 6.0 the worst grade)	504	1.80	.577	1	3.7
(Preliminary) University Grade Average	Grade average in German university system (1.0 being the best & 6.0 the worst grade)	489	1.94	.551	1	3.67
Field of Studies						
Engineering	Dummy (Yes=1, 0 otherwise)	504	.141	-	0	1
Mathematics, Informatics and Natural Sciences	Dummy (Yes=1, 0 otherwise)	504	.169	-	0	1
Law, Economics and Social Sciences	Dummy (Yes=1, 0 otherwise)	504	.373	-	0	1
Linguistic and Cultural Sciences	Dummy (Yes=1, 0 otherwise)	504	.284	-	0	1
Other	Dummy (Yes=1, 0 otherwise)	504	.034	-	0	1
Status of Home Institution						
Private Home Institution	Dummy (Yes=1, 0 otherwise)	501	.022	-	0	1
Public Home Institution	Dummy (Yes=1, 0 otherwise)	501	.978	-	0	1
Type of Home Institution						
University	Dummy (Yes=1, 0 otherwise)	500	.764	-	0	1
Technical University	Dummy (Yes=1, 0 otherwise)	500	.154	-	0	1
University of Applied Sciences	Dummy (Yes=1, 0 otherwise)	500	.082	-	0	1
Location of Home Institution (Federal State)						
Baden-Wuerttemberg	Dummy (Yes=1, 0 otherwise)	500	.428	-	0	1
Bavaria	Dummy (Yes=1, 0 otherwise)	500	.136	-	0	1
Berlin	Dummy (Yes=1, 0 otherwise)	500	.074	-	0	1
Brandenburg	Dummy (Yes=1, 0 otherwise)	500	.01	-	0	1
Bremen	Dummy (Yes=1, 0 otherwise)	500	.012	-	0	1
Hamburg	Dummy (Yes=1, 0 otherwise)	500	.004	-	0	1
Hesse	Dummy (Yes=1, 0 otherwise)	500	.03	-	0	1
Mecklenburg-Hither Pomerania	Dummy (Yes=1, 0 otherwise)	500	.002	-	0	1
North Rhine-Westphalia	Dummy (Yes=1, 0 otherwise)	500	.10	-	0	1
Lower Saxony	Dummy (Yes=1, 0 otherwise)	500	.058	-	0	1
Rhineland-Palatinate	Dummy (Yes=1, 0 otherwise)	500	.076	-	0	1
Saarland	Dummy (Yes=1, 0 otherwise)	500	.02	-	0	1
Saxony	Dummy (Yes=1, 0 otherwise)	500	.032	-	0	1
Saxony-Anhalt	Dummy (Yes=1, 0 otherwise)	500	.008	-	0	1
Schleswig-Holstein	Dummy (Yes=1, 0 otherwise)	500	.008	-	0	1
Other	Dummy (Yes=1, 0 otherwise)	500	.002	-	0	1
Home Institution in one of the New Eastern States?	Dummy (Yes=1, 0 otherwise)	500	.052	-	0	1

Reputation of Home Institution						
Home Institution Ranking 1	CHE-Ranking 2010, School Grade (1.0 being best & 6.0 worst grade) assigned by students for the perceived teaching situation	348	2.42	.514	1.5	4.1
Home Institution Ranking 2	CHE-Ranking 2010, Percentage of professors perceiving this institution as leading in the specific subject area (allowed to name up to 5 universities)	324	.211	.210	.004	.836
Home Institution THE Top100	Dummy (Yes, i.e. listed as Top European University in THE Ranking 10/11=1, 0 otherwise)	504	.375	-	0	1
Home Institution Ranking 3	THE Ranking 2010/11 Top European Universities, Overall Scores (higher scores= better reputation)	189	52.22	6.06	47	67
Home Institution Ranking 4	THE Ranking 2010/11 Top European Universities, Overall Position (higher position=lower reputation)	189	53.93	25.93	9	79
Academic Degree Pursued						
'Bachelor'	Dummy (Yes=1, 0 otherwise)	495	.149	-	0	1
'Diplom'	Dummy (Yes=1, 0 otherwise)	495	.570	-	0	1
'Examen'	Dummy (Yes=1, 0 otherwise)	495	.097	-	0	1
'Magister'	Dummy (Yes=1, 0 otherwise)	495	.184	-	0	1
Duration of Study						
Semester	# of semesters studied	504	4.39	1.69	2	14
Elementary Student	Dummy (semester 1-3 at time of application=1, 0 otherwise)	504	.190	-	0	1
Previous Work Experience (Apprenticeship)						
Completed Vocational Training	Dummy (Yes=1, 0 otherwise)	504	.056	-	0	1
Second-Chance Education[109]	Dummy (Yes=1, 0 otherwise)	504	.018	-	0	1
Work Experience Part-Time						
Number of Part-Time Jobs	Total # of part-time jobs the applicant has mentioned in CV	504	1.92	1.72	0	9
(Previous or Current) Part-Time Job only at University	Dummy (Yes=1, 0 otherwise)	503	.157	-	0	1
(Previous or Current) Part-Time Jobs both at & outside University	Dummy (Yes=1, 0 otherwise)	503	.193	-	0	1
(Previous or Current) Part-Time Job only outside University	Dummy (Yes=1, 0 otherwise)	503	.416	-	0	1
No Part-Time Job	Dummy (Yes=1, 0 otherwise)	503	.235	-	0	1
Number of Internships	# of internships mentioned (CV)	504	1.50	1.37	0	9
Duration of Internships	Cumulative duration of internships (in months)	503	3.13	3.87	0	33

[109] Second-chance education describes the fact that a student did not receive the eligibility of university admission directly after 13 years of schooling (the traditional way), but had left school earlier and had been working for several years (or had completed an apprenticeship for instance) before returning to school education as an adult.

Other Scholarships						
'BAföG'	Dummy (Yes=1, 0 otherwise)	504	.226	-	0	1
Previous Merit-Based Stipend	Dummy (Yes=1, 0 otherwise)	504	.147	-	0	1
'Studienstiftung des Dt. Volkes'	Dummy (Yes=1, 0 otherwise)	504	.060	-	0	1
Also Applied for Other Scholarships	Dummy (Yes=1, 0 otherwise)	504	.421	-	0	1
Extracurricular Activites						
Extracurricular Activities Mentioned	Dummy (Yes=1, 0 otherwise)	504	.843	-	0	1
Type of Extracurricular Activity (Multiple Choices possible)						
Youth/Sports/Recreation	Dummy (Yes=1, 0 otherwise)	504	.482	-	0	1
Social	Dummy (Yes=1, 0 otherwise)	504	.248	-	0	1
Arts & Culture	Dummy (Yes=1, 0 otherwise)	504	.228	-	0	1
Political	Dummy (Yes=1, 0 otherwise)	504	.177	-	0	1
Clerical	Dummy (Yes=1, 0 otherwise)	504	.143	-	0	1
Student Association/Faculty	Dummy (Yes=1, 0 otherwise)	504	.212	-	0	1
Student Representation (High School)	Dummy (Yes=1, 0 otherwise)	504	.145	-	0	1
Student Exchange	Dummy (Yes=1, 0 otherwise)	504	.081	-	0	1
Number of Extracurricular Activities (ECAs)						
One Type of ECAs	Dummy (Yes=1, 0 otherwise)	504	.290	-	0	1
More than one Type of ECAs	Dummy (Yes=1, 0 otherwise)	504	.516	-	0	1
No Extracurricular Activity	Dummy (Yes=1, 0 otherwise)	504	.194	-	0	1
Strength of Extracurricular Activities (ECAs), measured as combination of leadership positions & assumed effort						
High ECA	Dummy (Yes=1, 0 otherwise)	504	.065	-	0	1
Medium ECA	Dummy (Yes=1, 0 otherwise)	504	.274	-	0	1
Low ECA	Dummy (Yes=1, 0 otherwise)	504	.504	-	0	1
No ECA	Dummy (Yes=1, 0 otherwise)	504	.157	-	0	1
Letter of Reference						
Recommendation	On a scale from 0 to 10	477	9.25	.708	5	10
Recommending Person=Professor	Dummy (Yes=1, 0 otherwise)	504	.681	-	0	1
Good Relation to Recommending Person	Dummy (Yes=1, 0 otherwise)	475	.834	-	0	1
Language Proficiency						
Language Skills	On a scale from 0 to 100	491	88.61	9.35	40	100
TOEFL	Dummy (Yes=1, 0 otherwise)	504	.562	-	0	1
Project-Specific Statements						
Existing Contacts to Guest Institution	Dummy (Yes=1, 0 otherwise)	504	.452	-	0	1
Tuition Fees	Dummy (Yes=1, 0 otherwise)	504	.655	-	0	1
Amount of Tuition Fees	in Thousand US$	494	14.22	13.98	0	80
Participant in Organized Study Progr.	Dummy (Yes=1, 0 otherwise)	500	.278	-	0	1
Private Guest Institution	Dummy (Yes=1, 0 otherwise)	504	.276	-	0	1
Top50 Guest Institution	Dummy (THE-Subject-Ranking 1-50 =1, 0 otherwise)	504	.437	-	0	1
Top10 Guest Institution	Dummy (THE-Subject-Ranking 1-10 =1, 0 otherwise)	504	.181	-	0	1
Top5 Guest Institution	Dummy (THE-Subject-Ranking 1-5=1, 0 otherwise)	504	.067	-	0	1
Guest Institution in Canada	Dummy (Yes=1, 0 otherwise)	504	.206	-	0	1
Guest Institution in US-Midwest	Dummy (Yes=1, 0 otherwise)	504	.131	-	0	1
Guest Institution in US-Northeast	Dummy (Yes=1, 0 otherwise)	504	.236	-	0	1
Guest Institution in US-South	Dummy (Yes=1, 0 otherwise)	504	.177	-	0	1
Guest Institution in US-West	Dummy (Yes=1, 0 otherwise)	504	.250	-	0	1

Applicant Characteristics						
Gender	Dummy (Female=1, 0 otherwise)	504	.486	-	0	1
Glasses	Dummy (Yes=1, 0 otherwise)	499	.178	-	0	1
Born in Germany	Dummy (Yes=1, 0 otherwise)	504	.918	-	0	1
Age	In years at time of application	504	21.87	1.45	19	32
Parents Mentioned in CV	Dummy (Yes=1, 0 otherwise)	504	.335	-	0	1
Parents=Academics	Dummy (mentioned as academic in CV=1, 0 otherwise)	504	.236	-	0	1
Professional Aim=Research/Science	Dummy (Yes=1, 0 otherwise)	434	.143	-	0	1
Pre-Selection Evaluator Characteristics and Pre-Selection Situation						
Evaluator Gender	Dummy (Female=1, 0 otherwise)	504	.188	-	0	1
Evaluator Age	In years at time of selection	498	53.56	9.33	33	68
Applications/Evaluator	# of applications one evaluator needs to assess	504	23.73	10.37	2	40
Length of Application	# of pages of application	504	22.01	8.17	5	147
Number of Additional Certificates	# of additional, i.e. non-required, certificates handed in	504	2.78	4.08	0	63
Applicant-Pre-Selection-Evaluator Similarity						
Gender Similarity	Dummy (same gender=1)	504	.540	-	0	1
Regional Similarity	Dummy (study/teach in same federal state=1, 0 otherwise)	498	.052	-	0	1
Institutional Similarity	Dummy (study/teach at same type of institution=1, 0 otherwise)	498	.845	-	0	1
Field-of-Study Similarity	Dummy (study/teach in the same field=1, 0 otherwise)	504	.651	-	0	1
Evaluation Committee Characteristics						
Size of Evaluation Committee	# of evaluators in committee	254	4.35	.941	3	7
Fraction of Female Evaluators	# of female evaluators divided by # of all evaluators in committee	254	.260	.210	0	1
Mainly Male Evaluators	Dummy (Yes, i.e. >50% male evaluators=1, 0 otherwise)	254	.815	-	0	1
Average Evaluator Age	Average age of evaluators in committee	254	52.41	5.42	43.5	62
Dispersion Evaluator Age	Standard deviation of evaluator age in committee	254	16.02	6.25	4.24	27.48
Interview Framework						
Interview Duration	Duration of interview (in minutes)	254	13.89	2.61	3	21
Interview Position	(# of applicants the committee has already interviewed) – 1	254	6.52	3.95	1	16
Interview Time: 09-09:59 am	Dummy (Yes=1, 0 otherwise)	254	.134	-	0	1
Interview Time: 10-10:59 am	Dummy (Yes=1, 0 otherwise)	254	.150	-	0	1
Interview Time: 11-11:59 am	Dummy (Yes=1, 0 otherwise)	254	.157	-	0	1
Interview Time: 12-01:30 pm	Dummy (Yes=1, 0 otherwise)	254	.154	-	0	1
Interview Time: 02-02:59 pm	Dummy (Yes=1, 0 otherwise)	254	.146	-	0	1
Interview Time: 03-03:59 pm	Dummy (Yes=1, 0 otherwise)	254	.087	-	0	1
Interview Time: 04-04:59 pm	Dummy (Yes=1, 0 otherwise)	254	.083	-	0	1
Interview Time: 05-06:59 pm	Dummy (Yes=1, 0 otherwise)	254	.091	-	0	1
Year						
Selection Year	Dummy (1=2008, 0=2007)	504	.518	-	0	1

Table 6-7 Descriptive Statistics

7 ECONOMETRIC ANALYSIS OF SELECTION SUCCESS

Given that an applicant has decided to hand in an application, we are able to observe whether or not he or she was successful in selection and has been awarded a scholarship. With the help of all the information revealed during the application process, we are able to empirically link individual success (or failure) to specific applicant and rater characteristics as well as other extraneous factors and can subsequently identify determinants of scholarship awarding success. But in order to be able to empirically investigate scholarship selection success, we first need to elaborate how success in this case can be defined. In the present context, several possibilities for defining scholarship selection success are available. How success is defined in turn ultimately determines the applicable empirical specification and testing strategy. Therefore, it will be discussed in the following section which empirical specifications may be applied before the results of the most suitable alternative(s) will be presented in subsequent chapters.

7.1 METHODOLOGICAL NOTE: EMPIRICAL SPECIFICATION POSSIBILITIES

As has been mentioned before, every applicant is assigned a specific metric score on a scale from 0-100 (both in pre- and in final selection). Accordingly, success could be defined in terms of this metric: The higher the (pre-)selection score, the higher the probability of being awarded the scholarship (or being invited to an interview). Being able to observe individual selection success on such a quantitative scale, ordinary least squares (OLS) regression would be the appropriate testing methodology (e.g. Hair et al. 2010). For every one-unit increase in any of the independent variables, the respective change in the predicted score can be computed (holding every other influence constant) and significant determinants of high (or low) selection scores can be identified. The empirical specification of the OLS model would look as follows

$$Y(Score) = \beta_0 + \sum_{i=1}^{n} \beta_i \overrightarrow{X_i} + \varepsilon_i$$

However, one of the prerequisites for applying OLS regression is the metric character of the endogenous variable Y – the selection score ranging from 0 to 100 in this context. As OLS regression assumes a linear relationship, a one unit increase in this scale needs to reflect the same change in success probability at every position in the distribution. However, evaluators are completely aware of the fact that applicants who surpass the threshold of 80 points on this scale are invited to an interview or awarded the scholarship respectively. Therefore, the increase from 79 to 80 points is likely to have a

different impact on selection success than the increase from 50 to 51 (definitely not successful) or 90 to 91 (definitely invited or awarded) for instance. Consequently, the metric nature of the (pre-) selection scores needs to be scrutinized. From figure 7-1 it can be derived that evaluators in fact tend to assign selection scores around 80 more frequently than selection scores far below or far above this threshold.

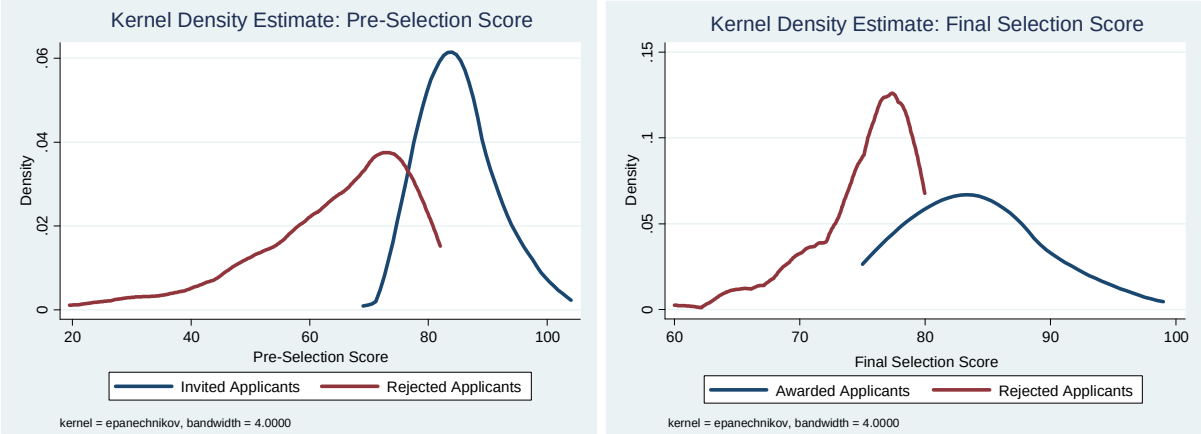

Figure 7-1 Kernel Density Estimates of Pre-Selection and Final Selection Scores

The accumulation of selection scores around this threshold could be accounted for by the use of the natural logarithm of selection scores as dependent variable, the application of quantile regression or a regression discontinuity (RD) design. Albeit, it is of major interest to this research who is awarded a scholarship (based on several signals and indices provided in an application). I.e. the decision whom to invite to an interview or whom to grant the stipend is a binary one – either somebody is accepted or rejected. Whether applicants just failed (i.e. they achieved scores ranging from 75 to 79 for instance) or whether they were clearly not suitable (reflected by scores far below 70) is only of minor importance to the present research. Hence, success (vs. failure) in scholarship awarding decisions should rather be defined in a binary way as 'awarded' vs. 'rejected'. In this case, the dependent variable is a qualitative one, as it can only attain two different values (accepted vs. rejected). This leads to a different empirical specification, as OLS regression techniques are only applicable to quantitative response variables (e.g. Wooldridge 2009).

For the analysis of dichotomous dependent variables, several possibilities exist (Gujarati and Porter 2009). Of these, the linear probability model (LPM) as well as probit and logit regression are the most frequently applied techniques. This is why their applicability in the present context will be discussed.

As the regressand in the current case is dichotomous and can only attain values of 0 (i.e. rejected) and 1 (accepted), typical linear regression models cannot directly be applied. In models where the regressand Y is qualitative, the objective is to determine the probability of an event to occur, i.e. Y attaining a certain value. In the case of dichotomous outcome variables – such as the scholarship awarding decision – the probability that the event will occur, i.e. $P(Y = 1)$, or will not occur ($P(Y = 0) = 1 - P(Y = 1)$) add up to 1. P is then modeled as a function of various explanatory variables X_i. The LPM directly models this conditional probability P and consequently its counterpart $1 - P$ as a linear function of the explanatory variables X_i (Gujarati and Porter 2009)

$$E(Y_i|X_i) = \beta_0 + \sum_{i=1}^{n} \beta_i \overrightarrow{X_i} + \varepsilon_i = P_i$$

The conditional expectation of this model $E(Y_i|X_i)$ is then interpreted as the conditional probability of $Y_i = 1$ which in turn can be determined by using simple OLS estimation (Gujarati and Porter 2009). The underlying assumption of LPM is consequently that the probability of Y_i equaling 1 increases linearly with X_i. However, as probabilities by definition need to range between 0 and 1, E needs to be restricted to this area. The difficulty to restrict the calculated values of LPM to values within this boundary is the most severe shortcoming of this model.[110] Usual OLS estimation does not take into account that $0 \leq E(Y_i|X_i) \leq 1$ as this is an inequality restriction and hence the calculated values $\widehat{Y_i}$ can attain values less than 0 and greater than 1 (Gujarati and Porter 2009).

In these cases, the calculated $\hat{Y}_i$ values need to be adapted in retrospect, i.e. $\hat{Y}_i$ is assumed to be 0 for negative model outcomes and $\hat{Y}_i$ is defined to be 1 for calculated $\hat{Y}_i$ values greater than 1. Hence, the LPM is easy to apply to categorical outcome variables as modeling and interpretation follow simple OLS rules, but does not specify the underlying model correctly and needs to be adapted in several cases. In order to avoid this retrospective adaptation, other regression models for categorical outcome variables a priori limit the values of $\hat{Y}_i$ to the range from 0 to 1. This improves model specification, but on the other hand exacerbates coefficient interpretation. Among these models, logit and probit regression will be presented here as these are regression methods which account for the aforementioned shortcoming and restrict the range of the predicted values of $\hat{Y}_i$ between 0 and 1.

[110] Further problems are e.g. the non-normality of disturbances, heteroscedastic variances of disturbances as well as ambiguous values of the R^2 measure (Gujarati and Porter 2009).

Logistic regression (or the logit model) models the probability of $Y = 1$, i.e. being awarded a scholarship, as a cumulative logistic distribution function as follows

$$P_i = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i \overrightarrow{X_i} + \varepsilon_i)}} = \frac{1}{1 + e^{-Z_i}} = \frac{e^{Z_i}}{1 + e^{Z_i}}$$

$$1 - P_i = \frac{1}{1 + e^{Z_i}}$$

$$\frac{P_i}{1 - P_i} = \frac{e^{Z_i}}{1 + e^{-Z_i}} = e^{Z_i}$$

$$L_i = ln\left(\frac{P_i}{1 - P_i}\right) = Z_i = \beta_0 + \sum_{i=1}^{n} \beta_i \overrightarrow{X_i} + \varepsilon_i$$

Note that in the logit model, the probability is a priori, i.e. by definition, restricted to remain within the range of 0 to 1 and approaches these boundaries asymptotically (Gujarati and Porter 2009). This specification represents a better model fit than LPM which models this relationship as being linear. However, for this specification, the usual OLS procedure is not applicable as P_i is now nonlinear in X and the β_i. Nevertheless, through the use of a (logistic) linking function logistic regression combines several advantages: The transformation of probabilities into log odds ($ln\left(\frac{P_i}{1-P_i}\right)$) restricts the values of P_i to the range from 0 to 1, but allows Z and hence the logit L_i to vary from $-\infty$ to $+\infty$ (e.g. Urban 1993; Hosmer and Lemeshow 2000; Pampel 2000; Menard 2002; Long and Freese 2006; Gujarati and Porter 2009). As a result of this logit transformation, L_i is linear in X_i, but as stated before, the probability P_i and the β_i are not. Consequently, interpretation of probability changes induced by a one-unit change in one of the regressors cannot be interpreted as straightforward as in OLS or LPM models. Due to the selected linking function, only the change in the log odds (=logits) occurs as a linear function of X_i. Nonetheless, interpretation can be facilitated through the use of odds ratios instead of probabilities (Pampel 2000). The odds are simply the probability of an event occurring divided by the probability of an event not occurring (Menard 2002):

$$\frac{P_i}{1 - P_i} = Odds$$

Odds ratios are then calculated by dividing the odds of one group, e.g. female, divided by the odds of the other group, e.g. male (Menard 2002). Odds ratios can attain values ranging from 0 to $+\infty$. Values less than 1 reflect a lower probability of the event occurring than the event not occurring and odds ratios > 1 represent a higher probability of occurring than non-occurring. If the probability for both groups is the same, i.e. 0.5, the odds ratio equals 1. Comparing both equations (Odds vs. logit), one can easily see

that the logit is simply the natural logarithm of the odds. Hence, logistic regression coefficients (the β_i) can be interpreted as the change in log odds as a response to a one-unit change in one of the independent variables. As standard OLS is not applicable to logits, parameter estimation is conducted using the maximum-likelihood (ML) method (Backhaus et al. 2011).

Analogous to logit regression, probit regression also models P_i to remain within the boundaries of 0 and 1 by the use of a linking function (Gujarati and Porter 2009). The main differences between these two approaches are the theoretical assumption as well as the underlying cumulative distribution function (CDF). Whereas logit regression transforms the probability of an event occurring into odds ratios and assumes that the natural log of these odds is linearly related to the regressors, probit regression models the binary outcome with the help of a latent variable I (Gujarati and Porter 2009). This latent variable can be interpreted as a utility index which is determined by one or more explanatory variables. The larger the respective value of I, the greater the probability of the event occurring, in our case the greater the probability of being awarded a scholarship. More importantly, it is assumed in probit regression that a critical or a threshold level of I (I^*) exists: Utility indices equal to or greater than I^* will lead to $Y = 1$ and those below I^* will be associated with $Y = 0$ (Gujarati and Porter 2009). Although both the index I and the threshold level I^* are unobservable, it is assumed in the probit model that the latent variable I is normally distributed. Hence, $P\,(I^* \leq I)$ can be computed from the standardized normal CDF.

$$P_i = P(Y = 1|X) = P(I^* \leq I) = P\left(Z_i \leq \beta_0 + \sum_{i=1}^{n} \beta_i \overrightarrow{X_i} + \varepsilon_i\right) = F\left(\beta_0 + \sum_{i=1}^{n} \beta_i \overrightarrow{X_i} + \varepsilon_i\right)$$

This directly reveals the second difference between logit and probit regression. Whereas logit regression assumes a logistic distribution of P_i, the basis of probit regression is the standardized normal CDF (Gujarati and Porter 2009). In practice, both models obtain quite similar results[111] and do only differ slightly at the tails of their distribution (see e.g. Gujarati and Porter 2009, 572 for an illustration). For the ease of interpretation through the use of odds ratios, the logistic regression has been chosen and logit results only will be presented subsequently.[112]

[111] Actually, multiplying the probit coefficient by 1.81 (or multiplying the logit coefficient by 0.55) yields the respective logit (probit) coefficient (Gujarati and Porter 2009, 571).

[112] For a comparison of coefficients, probit estimations of all models are available in Appendix 2.

7.2 Empirical Results: Logit Regression

In order to test the hypotheses developed in Chapter 5 of this thesis, several logit models with differing sets of independent variables have been estimated and will be discussed in this section. Furthermore, the models do not only vary in the number of regressors included in the estimation, but also the dependent variable Y can be modeled in different ways. Whereas section 7.2.1 models the success of all applicants to be awarded the scholarship without differentiating between pre-selection and final selection success, section 7.2.2 discusses success at each selection stage independently. In section 7.2.3, an estimation modeling both pre- and final selection success simultaneously will be presented and compared to the previous results. Due to the non-linear nature of logit regression, only main effects will be presented in the empirical models. Interaction effects have also been tested and the result of the most important ones will be displayed and interpreted in the discussion section (Chapter 7.3).

7.2.1 Overall Success

As has been mentioned before, the major interest of this research is to find out who among the applicant pool is awarded a scholarship and why. It has already been concluded in Chapter 6 that the applicant pool represents a specific, positively self-selected group of students that differs from the entire German student population in many aspects. But given this self-selected group of those who decided to apply, who is successful in selection and who is not? This can be demonstrated by empirically modeling the dependent variable $Y = 1$ whenever an applicant has been awarded the scholarship and $Y = 0$ if he or she was rejected. If we do not differentiate between rejection after pre- and rejection after final selection, we consider the selection process as a black box. In this case, the overall success of an applicant can be modeled as a function of different independent variables

$$L_i = ln\left(\frac{P_i}{1 - P_i}\right) = Z_i = \beta_0 + \sum_{i=1}^{n} \beta_i \overrightarrow{X_i} + \varepsilon_i$$

In line with Hypothesis 1, several ability signals are expected to influence evaluator outcomes. While Estimation I models the influence of the most credible ability signals only – i.e. previous academic achievement, measured in grades – on individual awarding success, Estimation IIa-c additionally include further ability signals which can reasonably be expected to have an effect on the selection success of an applicant.

$$L_{I_i}(Y_{Aw} = 1) = \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i + \varepsilon_i$$

$$\begin{aligned} L_{IIa_i}(Y_{Aw} = 1) = {} & \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i \\ & + \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i \\ & + \beta_6 WorkExperience_i + \beta_7 SponsorshipBaföG_i + \beta_8 Sponsorshipmeritbased_i \\ & + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i + \beta_{11} PrivateInstitution_i \\ & + \beta_{12} THETop10_i + \beta_{13} GuestInstitutionRegion_i + \beta_{14} Number\ of\ ECA_i \\ & + \beta_{15} LOR_i + \beta_{16} Control_i + \varepsilon_i \end{aligned}$$

$$\begin{aligned} L_{IIb_i}(Y_{Aw} = 1) = {} & \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i \\ & + \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i \\ & + \beta_6 WorkExperience_i + \beta_7 SponsorshipBaföG_i \\ & + \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i \\ & + \beta_{11} PrivateInstitution_i + \beta_{12} THETop10_i + \beta_{13} GuestInstitutionRegion_i \\ & + \beta_{14} Number\ of\ ECA_i + \beta_{15} LOR_i + \beta_{16} Control_i + \varepsilon_i \end{aligned}$$

Estimation IIa and IIb only vary slightly in terms of how previous sponsorship influence is included in the model: Whereas Estimation IIa models the influence of any previous merit-based stipend, Estimation IIb includes a Dummy-Variable for recipients of the 'Studienstiftung des Deutschen Volkes' stipend only. Estimation IIc resembles Estimation IIb very closely, but additionally models the influence of the applicants' field of study on their success rates.

$$\begin{aligned} L_{IIc_i}(Y_{Aw} = 1) = {} & \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i \\ & + \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i \\ & + \beta_6 WorkExperience_i + \beta_7 SponsorshipBaföG_i \\ & + \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i \\ & + \beta_{11} PrivateInstitution_i + \beta_{12} THETop10_i + \beta_{13} GuestInstitutionRegion_i \\ & + \beta_{14} Number\ of\ ECA_i + \beta_{15} LOR_i + \beta_{16} FieldofStudy_i + \beta_{17} Control_i + \varepsilon_i \end{aligned}$$

Estimation III further includes individual difference factors which are not expected to influence selection success. Note that based on the available literature on discrimination in hiring, further individual difference factors, e.g. physical attractiveness, IM tactics and personality, are likely to influence evaluator decisions and hence, these influences should be controlled for. However, not all of these influences have been recorded during the investigated selection process. Consequently, these influences cannot be tested directly and are thus incorporated in the disturbance term ε_i only.

$$L_{III_i}(Y_{Aw} = 1) = \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i + \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i + \beta_6 WorkExperience_i + \beta_7 SponsorshipBafögG_i + \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i + \beta_{11} PrivateInstitution + \beta_{12} THETop10 + \beta_{13} GuestInstitutionRegion + \beta_{14} Number\ of\ ECA + \beta_{15} LOR + \beta_{16} Gender_i + \beta_{17} Glasses_i + \beta_{18} PlaceofBirth + \beta_{19} Age + \beta_{20} Parents_i + \beta_{21} Control_i + \varepsilon_i$$

Logit regression coefficients for all of the above mentioned estimations are provided in table 7-1. All model specifications are subsequently compared using different Goodness-of-Fit (GoF) measures. These include the following: [113]

Percentage of cases correctly classified: Assuming that every record with a predicted probability greater than 0.5 leads to a predicted outcome of 1 and every record with a predicted probability less than 0.5 leads to a predicted outcome of 0, the model outcomes can be compared to the real outcomes and can hence be classified as correct or incorrect (e.g. Hosmer and Lemeshow 2000; Backhaus et al. 2011).

Pseudo R^2: As logit regression coefficients are calculated through the use of ML-estimations, no 'real' R^2 measure as it is known from OLS estimations exists (Hair et al. 2010). However, several pseudo R^2 measures have been developed which resemble the 'real' R^2 only in terms of their range being restricted from 0 to 1 (Urban 1993).

1) **McFaddens (adjusted) R^2:** Is based on the comparison of the log-likelihood values of the full and the null, i.e. intercept only, model (McFadden 1974) and is defined as

$$McFaddens\ R^2 = 1 - \frac{\mathrm{LL}_{Full}}{\mathrm{LL}_{Intercept}}$$

whereas: $LL_F = Log - Likelihood\ of\ the\ Full\ Model$

$LL_{Intercept} = Log - Likelihood\ of\ the\ Intercept\ Model$

McFaddens-R^2 is a relative GoF measure as it indicates the relative improvement of the current model in comparison to the intercept model (Backhaus et al. 2011).

The adjusted version penalizes models with too many predictors by including a measure of the number of predictors (K). Note that adjusted McFaddens-R^2 can take negative values (Gordon 2012).

[113] For a detailed explanation of all GoF measures used here, see e.g. Veall and Zimmermann (1996), Hosmer and Lemeshow (2000) or Long and Freese (2006). An overview of all these R^2 can also be found at http://www.ats.ucla.edu/stat/mult_pkg/faq/general/psuedo_rsquareds.htm.

$$McFaddens\ Adjusted\ R^2 = 1 - \frac{\mathrm{LL}_{Full} - K}{\mathrm{LL}_{Intercept}}$$

2) **Cox-Snell-R²:** This pseudo R² designed by Cox and Snell (1989) reflects the improvement of the full model over the intercept model as follows:

$$Cox - Snell - R^2 = 1 - \left(\frac{\mathrm{L}_{Intercept}}{\mathrm{L}_{Full}}\right)^{\frac{2}{n}}$$

Note that Cox and Snell's pseudo R² has a maximum value that is below 1 (Backhaus et al. 2011).

3) **Nagelkerke R²:** The Nagelkerke R² adjusts the R² of Cox and Snell (1989) in such a way that the possible value range is extended to 1 (Backhaus et al. 2011). This is done by dividing the Cox-Snell-R² by the maximum R², i.e. by $1 - \left(\mathrm{L}_{Intercept}\right)^{\frac{2}{n}}$:

$$Nagelkerke\ R^2 = \frac{1 - \left(\frac{\mathrm{L}_{Intercept}}{\mathrm{L}_{Full}}\right)^{\frac{2}{n}}}{1 - \left(\mathrm{L}_{Intercept}\right)^{\frac{2}{n}}}$$

4) **McKelvey & Zavoina R²:** This pseudo R² follows the structure of the 'usual' R² and calculates the quotient of the variance of the latent variable and the sum of the latent variable variance and the standard error variance (McKelvey and Zavoina 1975):

$$McKelvey\ \&\ Zavoina\ R^2 = \frac{Variance\ of\ y^*}{Variance\ of\ y^* + Variance\ of\ \varepsilon}$$

5) **(Adjusted) Count R²:** The Count R² measure (Long and Freese 2006) does not follow the usual R² logic, but simply divides the number of correctly classified cases (see above) by the number of total counts:

$$Count\ R^2 = \frac{\#\ of\ Correct\ Cases}{Total\ Count}$$

As even with random guessing, e.g. saying everybody was awarded the scholarship, one would already correctly classify 50% (given a normal distribution of both outcomes) of the cases. In order to control for this baseline prediction, the adjusted Count R² subtracts the count of the most frequent outcome l from both the correct and the total counts (Long and Freese 2006):

$$Count\ R^2 = \frac{\#\ of\ Correct\ Cases - l}{Total\ - l}$$

In addition to the pseudo R^2 measures, another GoF measure with a slightly different approach will be used for model comparison: The AIC(*n). The **Akaike Information Criterion (AIC)** is a measure of the relative goodness of fit of multivariate models which have been estimated by the use of the ML-method (Akaike 1973) and is calculated by

$$AIC = -2\,LL + 2P$$

where P is the number of parameters in the statistical model, and LL is the maximized value of the likelihood function for the estimated model (Long and Freese 2006). The AIC is not only a GoF measure, but can be used as a means of model selection: Among a set of different models, the model with the minimum AIC value is the preferred model (von Auer 2007). Unlike Likelihood-Ratio (LR)-comparisons, model selection with the help of the AIC can also be conducted for models that are not nested (Long and Freese 2006).

Based on all of the above mentioned GoF measures, the best model will be selected and discussed subsequently. As logistic regression coefficient interpretation is exacerbated as a consequence of the logistic linking function, also marginal effects and odds ratios will be presented, but for the preferred model only. However, the reader needs to keep in mind that marginal effects in non-linear models cannot be interpreted globally – as it would be the case in OLS regression –, but only hold true for a specific combination of all the other independent variables. Whenever marginal effects will be presented in this thesis, they represent the marginal effect of a change in the variable of interest for a standard applicant (unless otherwise specified). Standard applicants possess average values of all metric independent variables, e.g. grades or language proficiency. For dummy variables however, average values are not useful. Consequently, standard applicants belong to the respective dummy category which has occurred most frequently within the applicant pool. Standard applicants exhibit the combination of characteristics described in Appendices 3-5. For the ease of readability, the respective values of all other independent variables will also be presented alongside with the marginal effects.

Logistic Regression Coefficients[114]		Estimation Number				
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**I**	**IIa**	**IIb**	**IIc**	**III**
Academic Achievement	High School Grade Average	**-1.194*** **	**-1.043*** **	**-0.975*** **	**-0.943*** **	**-0.908*** **
	University Grade Average	**-1.790*** **	**-1.708*** **	**-1.654*** **	**-1.692*** **	**-1.592*** **
Home Institution Characteristics	*Reference Category: University*					
	University of Applied Sciences	-/-	-0.941	-0.855	-0.605	-0.673
	Technical University	-/-	-0.501	-0.583	-0.549	-0.507
Field of Study	*Reference Category: Law, Economics and Social Sciences*					
	Engineering	-/-	-/-	-/-	0.597	-/-
	Mathematics, Informatics & Natural Sciences	-/-	-/-	-/-	0.278	-/-
	Linguistic and Cultural Sciences	-/-	-/-	-/-	0.233	-/-
	Other Fields of Study	-/-	-/-	-/-	0.916	-/-
Language Proficiency	Language Skills	-/-	**0.031***	**0.031***	0.032	0.029
	TOEFL	-/-	**0.807*** **	**0.908*** **	**0.948*** **	**0.859*** **
Work Experience Part-Time	*Reference Category: No Part-Time Job*					
	at University	-/-	-0.148	0.156	0.205	0.186
	both at University & outside University	-/-	-0.007	0.054	0.103	0.147
	outside University	-/-	-0.131	-0.070	-0.018	-0.009
Recipient of other Sponsorships	'BAföG'	-/-	0.105	0.110	0.149	0.118
	Any other Merit-Based Stipend	-/-	0.229	-/-	-/-	-/-
	'Studienstiftung des Deutschen Volkes'	-/-	-/-	**1.144***	**1.166***	**1.164***
	Applied for other Scholarships	-/-	0.341	0.340	0.301	0.407
Project-Specific Statements	Tuition Fees (in Thousand US)	-/-	-0.005	-0.004	-0.002	-0.004
	Private Guest Institution	-/-	0.270	0.267	0.345	0.317
	Guest Institution=Top10 University	-/-	0.023	-0.062	-0.095	0.014
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*					
	Canada	-/-	0.098	0.142	0.136	0.235
	US-Midwest	-/-	0.466	0.430	0.383	0.513
	US-Northeast	-/-	-0.282	-0.312	-0.359	-0.285
	US-South	-/-	0.146	0.161	0.111	0.146
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*					
	One Type of ECAs	-/-	**1.023** **	**0.983** **	**0.949** **	**1.004** **
	More Types of ECAs	-/-	**0.967** **	**0.911** **	**0.935** **	**0.952** **
Letter of Reference	Recommendation	-/-	0.238	0.239	0.240	0.231
	Recommending Person=Professor	-/-	**0.964*** **	**0.942*** **	**0.951*** **	**1.014*** **
	Good Relation to Recommending Person	-/-	0.694	0.692	0.703	0.698
Applicant Characteristics	Gender	-/-	-/-	-/-	-/-	0.072
	Glasses	-/-	-/-	-/-	-/-	0.308
	Born in Germany	-/-	-/-	-/-	-/-	0.849
	Age	-/-	-/-	-/-	-/-	**-0.272***
	Parents=Academics	-/-	-/-	-/-	-/-	-0.163
Control Variables	*Semester*	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Participant in Organized Study-Abroad Program	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Selection Year	-/-	*incl.*	*incl.*	*incl.*	*incl.*
Constant		**-1.092*** **	**-4.096*** **	**-4.128*** **	**-4.433*** **	**-5.151*** **
Observations		429	429	429	429	429
Pseudo R^2		0.193	0.286	0.293	0.297	0.306
*significant at * p <0.10 ** p <0.05 *** p <0.01*						

Table 7-1 Logistic Regression Coefficients Estimations I-III

[114] Pairwise correlations between all independent variables have been calculated for all models presented below. The results of these multicollinearity tests are available on request.

Goodness of Fit and Model Comparison	Estimation Number[115] I	IIa	IIb	IIc	III
Cases correctly classified	74.59%	77.86%	78.09%	**78.55%**	77.86%
Log-Likelihood Full Model	-215.007	-190.239	-188.420	-187.229	**-184.844**
LR	102.779	152.316	155.954	158.335	**163.106**
(Prob>LR)	(0.000)	(0.000)	(0.000)	(0.000)	**(0.000)**
Mc Fadden's (Adjusted) R^2	0.193	0.286	0.293	0.297	**0.306**
	(0.182)	(0.177)	**(0.184)**	(0.173)	(0.179)
Cox-Snell R^2	0.213	0.299	0.305	0.309	**0.316**
Nagelkerke R^2	0.300	0.420	0.429	0.434	**0.445**
McKelvey and Zavoina's R^2	0.361	0.517	0.522	0.528	**0.542**
(Adjusted) Count R^2	0.746	0.779	0.781	0.779	**0.786**
	(0.187)	(0.291)	(0.299)	(0.291)	**(0.313)**
AIC	1.016	1.022	**1.014**	1.027	1.020
*AIC*n*	436.015	438.477	**434.839**	440.458	437.688

Table 7-2 Goodness-of-Fit Measures Estimations I-III

As can be seen in table 7-2, all of the presented models do explain stipend awarding decisions better than the intercept (or null) model, as all Likelihood-Ratio (LR) tests lead to the rejection of the null-hypothesis.[116] Additionally, the pseudo R^2 values are comparably high for logit models.[117] Following Estimation IIc, one would correctly classify 78.55% of all cases, but adjusted for the baseline correct classification rate (adjusted Count R^2), Estimation III provides the best classification. Additionally, Estimation III attains the highest values in most of the pseudo R^2.

As Estimation III provides the best fit in most of the GoF measures[118], the results from this estimation will be discussed subsequently. Considering the AIC however, one would prefer Estimation IIb, but as all AIC values are quite comparable, Estimation III was chosen due to the relatively high pseudo R^2 values. Consequently, results from this estimation will be discussed in detail and marginal effects for a standard applicant derived from Estimation III results are provided in table 7-3.[119]

[115] Note that Estimation IIa and Estimation IIc are not nested in Estimation III.
[116] In an LR-test, the null-hypothesis states that all of the coefficients derived from the model are equal to zero, i.e. they do not have an influence on the model outcome (Long and Freese 2006).
[117] Usually, in logit regression, low pseudo R^2 values are the norm (Hosmer and Lemeshow 2000) and values of $0.2 < R^2 < 0.4$ are already considered excellent results (Urban 1993).
[118] The respective best GoF measure is always printed in bold in all of the GoF-tables.
[119] As a consequence of missing values in important explanatory variables some cases had to be excluded from the analysis. However, the sample of applicants finally included in the subsequent estimations does only differ slightly from the entire applicant pool as can be derived from the descriptive statistics in Appendix 6.

Marginal Effects after Estimation III		dy/dx	Value of X
Academic Achievement	High School Grade Average	**-.1552****	1.76
	University Grade Average	**-.2723*****	1.89
Home Institution Characteristics	*Reference Category: University*		
	University of Applied Sciences[i]	-.0939	0
	Technical University[i]	-.0745	0
Language Proficiency	Language Skills	.0050	88.85
	TOEFL[i]	**.1128***	1
Work Experience Part-Time	*Reference Category: No Part-Time Job*		
	at University[i]	.0334	0
	both at University & outside University[i]	.0261	0
	outside University [i]	-.0015	1
Recipient of other Sponsorships	'BAföG'[i]	.0209	0
	'Studienstiftung des Deutschen Volkes' [i]	.2542	0
	Applied for other Scholarships[i]	.0774	0
Project-Specific Statements	Tuition Fees (in Thousand US)	-.0007	14.643
	Private Guest Institution[i]	.0590	0
	Guest Institution=Top10 University	-.0024	0
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*		
	Canada[i]	.0428	0
	US-Midwest[i]	.0999	0
	US-Northeast[i]	-.0449	0
	US-South[i]	.0259	0
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*		
	One Type of ECAs[i]	**.2144***	0
	More Types of ECAs[i]	**.1213***	1
Letter of Reference	Recommendation	.0396	9.26
	Recommending Person=Professor[i]	**.1267***	1
	Good Relation to Recommending Person[i]	.0965	1
Applicant Characteristics	Gender[i]	.0125	0
	Glasses[i]	.0572	0
	Born in Germany[i]	.1118	1
	Age	**-.0465***	21.84
	Parents=Academics[i]	-.0266	0
Control Variables	*Semester*	incl.	4.39
	Existing Contacts to Guest Institution	incl.	0
	Participant in Organized Study-Abroad Program	incl.	0
	Selection Year	incl.	1
[i] *dy/dx is for discrete change of dummy variable from 0 to 1*			

Table 7-3 Marginal Effects for Standard Applicant after Estimation III

As expected in Hypothesis 2, both high school and university grades have an important influence on selection success. Applicants with a high school grade average of 'only' 2.76 c.p. have a 15.5 percentage points (pps) lower probability of being awarded the scholarship than standard applicants. Poor university grades are even punished more severely, as a candidate possessing all characteristics of a standard applicant, but differing only in terms of having achieved 'only' a university grade average of 2.89 has a 27.3 pps lower chance of being awarded the stipend. Other effective ability signals are the provision of a TOEFL (+0.1128) as well as stating one or more types of ECAs in the CV. Here it seems particularly advantageous to concentrate on one type of activity only,

as this leads to a 21.4 pps higher chance of being successful with the application, whereas stating more than one type of activities is also rewarded (in comparison to stating no ECA at all), but for a standard applicant (on all other dimensions), the increase in 'winning probability' is only 12.1 pps. Finally, providing an LOR issued by a professor (instead of a research assistant or assistant professor) is also considered an effective signal (+0.127) although the content, i.e. the recommendation itself, does not significantly increase awarding probabilities. All other theoretically expected ability signals such as type of home institution, language skills, (part-time) work experience, previous sponsorships/awards or guest institution characteristics do not have an effect on award probability.[120] In line with the hypotheses, individual difference factors in general do not affect awarding decisions. Only one of the individual difference factors that could be empirically tested has an influence on selection success: Age. An applicant being standard on all other dimensions, but aged 23 at the time of application, i.e. one year older than the standard applicant, has a 4.7 pps lower chance of being awarded the stipend.

Although Estimation III delivers a quite satisfactory model fit, it does not account for the peculiarities of each selection stage. In the overall success measure, it was only observed whether or not somebody was accepted or rejected in the end, but it was not distinguished between somebody who was rejected in pre-selection and somebody who 'survived' the first selection round, but was rejected after the interview. As the literature review has revealed that determinants of selection success vary widely between pre- and final selection, it is worthwhile looking at both stages separately.

[120] At least this holds true for a standard applicant. Due to the non-linear nature of logistic regression, it might be that the aforementioned ability signals are effective for candidates with a different combination of some or all the other independent variables.

7.2.2 Separate Investigation of Each Selection Stage

The literature review has revealed that recruiters (or more general evaluators) base their decisions in paper-based pre-selection on different criteria than in person-to-person interviews. Grades are e.g. expected to play a more important role in paper-based selection than in final selection (compare Hypothesis 3). In the case of stipend awarding decisions, the same result might be found. When looking at the present selection process, a comparison of both pre- and final selection scores assigned by evaluators is especially suited to figure out whether or not evaluators base their decision on the same or different evaluation criteria. If we assume that the same signals are effective in both pre- and final selection, pre- and final selection scores are expected to be highly correlated, meaning that somebody who was able to achieve a high score in pre-selection will be likely to also achieve a high score in final selection. Figure 7-2 shows the distribution of pre-selection scores (abscissa) and final selection scores (ordinate) for all applicants who have been interviewed in the final round.[121]

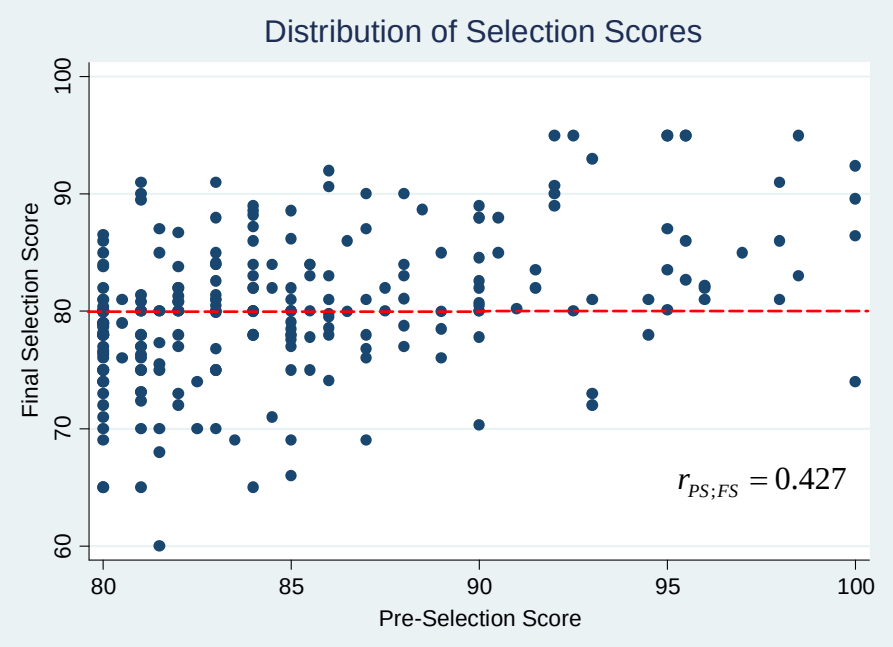

Figure 7-2 Distribution of Pre- and Final Selection Scores in Comparison

It appears from figure 7-2 that there is a positive correlation between pre- and final selection scores. However, this correlation is not at all perfect, as $r_{PS;FS} = 0.427$ ($p < 0.001$) and a high pre-selection score (>90) does not guarantee final selection success, i.e. being awarded a scholarship, and vice versa.[122] Hence, a separate investigation of each stage is warranted.

[121] Applicants who have been rejected after pre-selection were not assigned a final selection score.
[122] All applicants with final selection scores ≥ 80 are awarded the stipend (red line in figure 7-2).

7.2.2.1 Pre-Selection Success

In order to compare the separate investigation of pre-selection success with the overall success probabilities, the same regressors as in Estimation I-III have been chosen for Estimations IV-VI. This time however, the regressand Y is different. The dependent variable is no longer Y_{Aw} (yes or no), but Y_{Inv}. This variable equals 1 whenever an applicant was successful in pre-selection, i.e. invited to an interview, and 0 whenever an applicant was rejected during pre-selection.

$$L_{IV\,i}(Y_{Inv}=1)=\beta_0+\beta_1 AcademicAchievement1_i+\beta_2 AcademicAchievement2_i+\varepsilon_i$$

$$\begin{aligned}L_{Va_i}(Y_{Inv}=1)=\;&\beta_0+\beta_1 AcademicAchievement1_i+\beta_2 AcademicAchievement2_i\\&+\beta_3 HomeInstitutionChar_i+\beta_4 LanguageSkills_i+\beta_5 TOEFL_i\\&+\beta_6 WorkExperience_i+\beta_7 SponsorshipBaf\ddot{o}G_i+\beta_8 Sponsorshipmeritbased_i\\&+\beta_9 Sponsorshipapplication_i+\beta_{10} Tuition_i+\beta_{11} PrivateInstitution_i\\&+\beta_{12} THETop10_i+\beta_{13} GuestInstitutionRegion_i+\beta_{14} Number\ of\ ECA_i\\&+\beta_{15} LOR_i+\beta_{16} Control_i+\varepsilon_i\end{aligned}$$

$$\begin{aligned}L_{Vb_i}(Y_{Inv}=1)=\;&\beta_0+\beta_1 AcademicAchievement1_i+\beta_2 AcademicAchievement2_i\\&+\beta_3 HomeInstitutionChar_i+\beta_4 LanguageSkills_i+\beta_5 TOEFL_i\\&+\beta_6 WorkExperience_i+\beta_7 SponsorshipBaf\ddot{o}G_i\\&+\beta_8 SponsorshipStudienstiftung_i+\beta_9 Sponsorshipapplication_i+\beta_{10} Tuition_i\\&+\beta_{11} PrivateInstitution_i+\beta_{12} THETop10_i+\beta_{13} GuestInstitutionRegion_i\\&+\beta_{14} Number\ of\ ECA_i+\beta_{15} LOR_i+\beta_{16} Control_i+\varepsilon_i\end{aligned}$$

$$\begin{aligned}L_{VI\,i}(Y_{Inv}=1)=\;&\beta_0+\beta_1 AcademicAchievement1_i+\beta_2 AcademicAchievement2_i\\&+\beta_3 HomeInstitutionChar_i+\beta_4 LanguageSkills_i+\beta_5 TOEFL_i\\&+\beta_6 WorkExperience_i+\beta_7 SponsorshipBaf\ddot{o}G_i\\&+\beta_8 SponsorshipStudienstiftung_i+\beta_9 Sponsorshipapplication_i+\beta_{10} Tuition_i\\&+\beta_{11} PrivateInstitution+\beta_{12} THETop10+\beta_{13} GuestInstitutionRegion\\&+\beta_{14} Number\ of\ ECA+\beta_{15} LOR+\beta_{16} Gender_i+\beta_{17} Glasses_i\\&+\beta_{18} PlaceofBirth+\beta_{19} Age+\beta_{20} Parents_i+\beta_{21} Control+\varepsilon_i\end{aligned}$$

In addition to the independent variables that have been tested in Estimations I-III, it is now also possible to include pre-selection specific (situational) variables such as evaluator gender and age in the model. This is done in Estimation VIIa and VIIb. Analogous to Estimation IIc, Estimation VIIb additionally includes the applicant's field of study.

$$L_{VIIa_i}(Y_{Inv} = 1) = \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i + \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i + \beta_6 WorkExperience_i + \beta_7 SponsorshipBafög_i + \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i + \beta_{11} PrivateInstitution + \beta_{12} THETop10 + \beta_{13} GuestInstitutionRegion + \beta_{14} Number\ of\ ECA + \beta_{15} LOR + \beta_{16} Gender_i + \beta_{17} Glasses_i + \beta_{18} PlaceofBirth + \beta_{19} Age + \beta_{20} Parents_i + \beta_{21} EvaluatorGender_i + \beta_{22} EvaluatorAge_i + \beta_{23} Control + \varepsilon_i$$

$$L_{VIIb_i}(Y_{Inv} = 1) = \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i + \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i + \beta_6 WorkExperience_i + \beta_7 SponsorshipBafög_i + \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i + \beta_{11} PrivateInstitution + \beta_{12} THETop10 + \beta_{13} GuestInstitutionRegion + \beta_{14} Number\ of\ ECA + \beta_{15} LOR + \beta_{16} Gender_i + \beta_{17} Glasses_i + \beta_{18} PlaceofBirth + \beta_{19} Age + \beta_{20} Parents_i + \beta_{21} EvaluatorGender_i + \beta_{22} EvaluatorAge_i + \beta_{23} FieldofStudy_i + \beta_{24} Control + \varepsilon_i$$

The logistic regression coefficients obtained from model Estimations IV to VIIb are displayed in table 7-4. Analogous to the previous section, model fit will be analyzed and compared with the aforementioned GoF measures in table 7-5. Marginal effects for a standard applicant will be presented in table 7-6 for the most suitable estimation only.

Logistic Regression Coefficients		Estimation Number					
Dependent Variable: Invitation (Yes=1 No=0)		**IV**	**Va**	**Vb**	**VI**	**VIIa**	**VIIb**
Academic Achievement	High School Grade Average	**-1.402*** **	**-1.195*** **	**-1.244*** **	**-1.257*** **	**-1.434*** **	**-1.679*** **
	University Grade Average	**-1.649*** **	**-1.531*** **	**-1.471*** **	**-1.561*** **	**-1.579*** **	**-1.726*** **
Home Institution Characteristics	*Reference Category: University*						
	University of Applied Sciences	-/-	**-1.568** **	**-1.549** **	**-1.288***	-0.797	-0.287
	Technical University	-/-	-0.540	-0.575	-0.453	-0.143	-0.134
Field of Study	*Reference Category: Law, Economics and Social Sciences*						
	Engineering	-/-	-/-	-/-	-/-	-/-	**1.180***
	Mathematics, Informatics & Natural Sciences	-/-	-/-	-/-	-/-	-/-	-0.202
	Linguistic and Cultural Sciences	-/-	-/-	-/-	-/-	-/-	**0.970** **
	Other Fields of Study	-/-	-/-	-/-	-/-	-/-	0.631
Language Proficiency	Language Skills	-/-	**0.053*** **	**0.052*** **	**0.050** **	**0.046** **	**0.046** **
	TOEFL	-/-	**1.454*** **	**1.551*** **	**1.492*** **	**1.483*** **	**1.579*** **
Work Experience Part-Time	*Reference Category: No Part-Time Job*						
	at University	-/-	-0.234	-0.212	-0.134	-0.247	-0.289
	both at University & outside University	-/-	**0.928***	**0.994** **	**1.077** **	**1.209** **	**1.238** **
	outside University	-/-	-0.254	-0.178	-0.203	-0.338	-0.255
Recipient of other Sponsorships	'BAföG'	-/-	-0.361	-0.341	-0.421	-0.553	-0.509
	Any other Merit-Based Stipend	-/-	**0.897** **	-/-	-/-	-/-	-/-
	'Studienstiftung des Dt. Volkes'	-/-	-/-	**2.337** **	**2.306** **	**2.160***	**2.280***
	Applied for other Scholarships	-/-	-0.243	-0.274	-0.196	-0.065	-0.184
Project-Specific Statements	Tuition Fees (in Thousand US)	-/-	**0.018***	**0.020***	0.018	0.015	0.021
	Private Guest Institution	-/-	**-0.720***	-0.715	-0.711	**-0.912***	**-0.870***
	Guest Institution=Top10 University	-/-	**-0.753***	**-0.823** **	**-0.740***	**-0.724***	-0.593
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*						
	Canada	-/-	0.016	0.091	0.043	0.202	0.367
	US-Midwest	-/-	-0.020	0.029	0.049	0.329	0.293
	US-Northeast	-/-	**0.873***	**0.906***	**0.985***	**1.205** **	**1.253** **
	US-South	-/-	0.403	0.488	0.497	0.577	0.567
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*						
	One Type of ECAs	-/-	**0.895** **	**0.945** **	**0.835***	**1.008** **	**1.021** **
	More Types of ECAs	-/-	**1.108*** **	**1.114*** **	**1.107** **	**1.303*** **	**1.330*** **
Letter of Reference	Recommendation	-/-	**0.629** **	**0.651** **	**0.612** **	**0.698** **	**0.726*** **
	Recommending Person=Professor	-/-	**1.076*** **	**1.089*** **	**1.261*** **	**1.353*** **	**1.582*** **
	Good Relation to Recommender	-/-	**1.641*** **	**1.718*** **	**1.783*** **	**1.835*** **	**1.755*** **
Applicant Characteristics	Gender	-/-	-/-	-/-	0.300	0.201	0.188
	Glasses	-/-	-/-	-/-	0.323	0.363	0.366
	Born in Germany	-/-	-/-	-/-	**1.655** **	**1.768*** **	**1.719*** **
	Age	-/-	-/-	-/-	**-0.358** **	**-0.403** **	**-0.399** **
	Parents=Academics	-/-	-/-	-/-	**-0.677***	**-0.846** **	**-0.894** **
Rater Characteristics	Evaluator Gender	-/-	-/-	-/-	-/-	**1.435*** **	**1.600*** **
	Evaluator Age	-/-	-/-	-/-	-/-	**0.043** **	**0.039***
Control Variables	*Semester*	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
	Part. in Organized Study-Abroad Program	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
	Selection Year	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
Constant		**0.346** **	**-3.611*** **	**-3.780*** **	**-5.485*** **	**-6.259*** **	**-6.911*** **
Observations		423	423	423	423	423	423
Pseudo R^2		0.226	0.425	0.430	0.463	0.481	0.497
*significant at * p <0.10 ** p <0.05 *** p <0.01*							

Table 7-4 Logistic Regression Coefficients Estimations IV-VIIb

Goodness of Fit and Model Comparison	Estimation Number[123] IV	Va	Vb	VI	VIIa	VIIb
Cases correctly classified	72.01%	82.27%	82.51%	84.87%	**85.58%**	85.11%
Log-Likelihood Full Model	-223.530	-165.935	-164.535	-155.144	-149.989	**-145.210**
LR	130.515	245.706	248.506	267.287	277.597	**287.155**
(Prob>LR)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	**(0.000)**
Mc Fadden's (Adjusted) R²	0.226	0.425	0.430	0.463	0.481	**0.497**
	(0.216)	(0.325)	(0.330)	(0.345)	(0.356)	**(0.359)**
Cox-Snell R²	0.265	0.441	0.444	0.468	0.481	**0.493**
Nagelkerke R²	0.356	0.592	0.597	0.629	0.646	**0.662**
McKelvey and Zavoina's R²	0.371	0.656	0.671	0.710	0.737	**0.756**
(Adjusted) Count R²	0.721	0.823	0.825	0.849	**0.856**	0.851
	(0.348)	(0.586)	(0.591)	(0.646)	**(0.663)**	(0.652)
AIC	1.071	0.922	0.915	0.894	0.879	**0.876**
AIC*n	453.060	389.869	387.069	378.288	371.978	**370.420**

Table 7-5 Goodness-of-Fit Measures Estimations IV-VIIb

Again, all presented models explain invitation decisions better than the intercept (or null) model, as all Likelihood-Ratio (LR) tests lead to the rejection of the null-hypothesis. Additionally, the pseudo R^2 values are extraordinarily high not only for logit models in general, but also in comparison to the values obtained in Estimations I to III. Following Estimation VIIa, one would correctly classify 85.58% of all cases. Adjusted for the baseline correct classification rate (adjusted Count R^2), 66.3% of cases would be correctly classified. Here, Estimation VIIb has the highest values in most of the R^2, except for the (adjusted) Count R^2. Additionally, the AIC value is minimal in Estimation VIIb which would prompt the reader to prefer this estimation to all other models. However, taking a look back at the different empirical specifications, it becomes obvious that Estimation VIIb exceptionally models the field of study. As logistic regression coefficients do not only depend on the respective value of the independent variable at stake, but also differ according to the values the other independent variables attain, one needs to consider that Estimation VIIb only models the respective influences for an applicant studying Law, Economics or Social Sciences (reference category). Consequently, all Estimation VIIb coefficients model the influence for this specific group of students only. In order to avoid such a sample restriction, the coefficients obtained from Estimation VIIa – the second best choice in all of the other GoF measures – will be discussed instead.

[123] Please note that Estimation Va is not nested in Estimation VIIb.

Marginal Effects after Estimation VIIa		dy/dx	Value of X
Academic Achievement	High School Grade Average	**-.2516****	1.76
	University Grade Average	**-.2770****	1.89
Home Institution Characteristics	*Reference Category: University*		
	University of Applied Sciences[i]	-.1676	0
	Technical University[i]	-.0261	0
Language Proficiency	Language Skills	**.0081***	88.76
	TOEFL[i]	**.1128***	1
Work Experience Part-Time	*Reference Category: No Part-Time Job*		
	at University[i]	-.0462	0
	both at University & outside University[i]	**.1464***	0
	outside University [i]	-.0539	1
Recipient of other Sponsorships	'BAföG'[i]	-.1109	0
	'Studienstiftung des Deutschen Volkes' [i]	**.1943****	0
	Applied for other Scholarships[i]	-.0116	0
Project-Specific Statements	Tuition Fees (in Thousand US)	.0026	14.652
	Private Guest Institution[i]	-.1954	0
	Guest Institution=Top10 University	-.1502	0
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*		
	Canada[i]	.0335	0
	US-Midwest[i]	.0524	0
	US-Northeast[i]	**.1461***	0
	US-South[i]	.0854	0
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*		
	One Type of ECAs[i]	**.1302***	0
	More Types of ECAs[i]	**.2923****	1
Letter of Reference	Recommendation	**.1224***	9.27
	Recommending Person=Professor[i]	**.3049*****	1
	Good Relation to Recommending Person[i]	**.4209*****	1
Applicant Characteristics	Gender[i]	.0333	0
	Glasses[i]	.0574	0
	Born in Germany[i]	**.4054*****	1
	Age	**-.0707***	21.85
	Parents=Academics[i]	**-.1792***	0
Evaluator Characteristics	Evaluator Gender[i]	**.1616***	0
	Evaluator Age	.0075	53.61
Control Variables	*Semester*	incl.	4.39
	Existing Contacts to Guest Institution	incl.	0
	Participant in Organized Study-Abroad Program	incl.	0
	Selection Year	incl.	1
[i] *dy/dx is for discrete change of dummy variable from 0 to 1*			

Table 7-6 Marginal Effects for Standard Applicant after Estimation VIIa

As can be derived from the GoF-comparison, the amount of explained variance is greater in pre- than in overall selection success. Many of the theoretically expected signals have a statistically significant impact on pre-selection success. For a standard applicant, the most effective among these are high school and university grades, previous part-time jobs both at the university and with an external employer, LORs from a professor who indicates to know the applicant well and mentioning more than one type of ECAs.

In pre-selection, poor grades (both in university and high school) are penalized severely. A standard applicant has a 27.7 (25.2) pps higher probability of being invited to a selection interview than an applicant who has only achieved university (high school) grades averaging 2.89 (2.76). When it comes to language proficiency, an additional point on the language-skills scale leads to a 0.8 pps increase in invitation probability. Again, applicants handing in a TOEFL have an increased chance of being invited to a final selection interview (+0.113). In terms of previous (part-time) work experience, applicants indicating they already had part-time jobs both at the university and with an external employer have a 14.6 pps higher probability of being invited than an applicant without any part-time job. Applicants who have already been awarded one of the most renowned German stipends and are hence sponsored by the 'Studienstiftung des Deutschen Volkes' c.p. have an increased probability of 'surviving' the pre-selection round (+0.194). Receiving a means-tested scholarship however has no significant (positive or negative) impact on pre-selection success. Again, indicating to be active in one or more ECAs increases invitation probabilities. In paper-based pre-selection, however, it appears to be especially advantageous to state more than one type of ECAs. This leads to a 0.292 increase in invitation probabilities whereas stating to be active in only one type of ECAs 'only' increases the probability of pre-selection success by 0.13. In line with the expectations, LORs are mainly effective in paper-based pre-selection. For a standard applicant, an increase of one unit on the recommendation scale leads to a 12.2 pps increase in invitation probability. However, it is again more important who issued the LOR: Whenever the recommendation is made by a full professor, pre-selection success probabilities increase by 0.305. Moreover, if the recommender states to know the applicant well, the applicant has an increased chance of being invited to an interview (+0.42) compared to someone whose recommender does not confirm a good relation to the applicant. Assuming a rational decision process, individual difference factors were not hypothesized to influence evaluator decisions. However, several of these indices do have an impact on invitation probabilities. Whereas neither applicants of a specific gender nor those wearing glasses are discriminated against, other individual difference factors influence pre-selection outcomes. Applicants born in Germany have a 40.5 pps higher probability of being invited to an interview than those born outside of Germany. Younger applicants and applicants not stating their parents as being academics are preferred over older ones (-0.071) and those indicating their academic background (-0.179). Finally, also evaluator individual difference factors have been found to influence pre-selection success: Whenever standard applicants are assessed by a female evaluator, their pre-selection success probability increases by 0.162.

7.2.2.2 Final Selection Success

Whenever applicants have managed to 'survive' the first selection round, they are interviewed by a selection committee. Final selection success then can be measured as the probability of being awarded the scholarship, conditional on being invited to an interview (and actually showing up)[124].

Estimations VIII-Xb model this conditional probability of $Y_{Aw} = 1|Y_{Inv} = 1$ by using the same independent variables that have already been included in the estimations of overall and pre-selection success.

$$L_{VIII_i}(Y_{Aw} = 1|Y_{Inv} = 1) = \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i + \varepsilon_i$$

$$\begin{aligned} L_{IXa_i}(Y_{Aw} = 1|Y_{Inv} = 1) \\ &= \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i \\ &+ \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i \\ &+ \beta_6 WorkExperience_i + \beta_7 SponsorshipBaföG_i + \beta_8 Sponsorshipmeritbased_i \\ &+ \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i + \beta_{11} PrivateInstitution_i \\ &+ \beta_{12} THETop10_i + \beta_{13} GuestInstitutionRegion_i + \beta_{14} Number\ of\ ECA_i \\ &+ \beta_{14} LOR_i + \beta_{15} Control_i + \varepsilon_i \end{aligned}$$

$$\begin{aligned} L_{IXb_i}(Y_{Aw} = 1|Y_{Inv} = 1) \\ &= \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i \\ &+ \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i \\ &+ \beta_6 WorkExperience_i + \beta_7 SponsorshipBaföG_i \\ &+ \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i \\ &+ \beta_{11} PrivateInstitution_i + \beta_{12} THETop10_i + \beta_{13} GuestInstitutionRegion_i \\ &+ \beta_{14} Number\ of\ ECA_i + \beta_{14} LOR_i + \beta_{15} Control_i + \varepsilon_i \end{aligned}$$

[124] Please remember that only 254 of the 272 invited applicants attended the interview.

$$L_{Xa_i}(Y_{Aw} = 1|Y_{Inv} = 1)$$
$$= \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i$$
$$+ \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i$$
$$+ \beta_6 WorkExperience_i + \beta_7 SponsorshipBafög_i$$
$$+ \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i$$
$$+ \beta_{11} PrivateInstitution + \beta_{12} THETop10 + \beta_{13} GuestInstitutionRegion$$
$$+ \beta_{14} Number\ of\ ECA + \beta_{14} LOR + \beta_{15} Control + \beta_{16} Gender_i + \beta_{17} Glasses_i$$
$$+ \beta_{18} PlaceofBirth + \beta_{19} Age + \beta_{20} Parents_i + \varepsilon_i$$

$$L_{Xb_i}(Y_{Aw} = 1|Y_{Inv} = 1)$$
$$= \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i$$
$$+ \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i$$
$$+ \beta_6 WorkExperience_i + \beta_7 SponsorshipBafög_i$$
$$+ \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i$$
$$+ \beta_{11} PrivateInstitution + \beta_{12} THETop10 + \beta_{13} GuestInstitutionRegion$$
$$+ \beta_{14} Number\ of\ ECA + \beta_{14} LOR + \beta_{15} Control + \beta_{16} Gender_i + \beta_{17} Glasses_i$$
$$+ \beta_{18} PlaceofBirth + \beta_{19} Age + \beta_{20} Parents_i + \beta_{21} FieldofStudy_i + \varepsilon_i$$

Logistic regression results of all of the above specified estimations are displayed in table 7-7. GoF measures are provided subsequently in table 7-8.

Logistic Regression Coefficients		Estimation Number				
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**VIII**	**IXa**	**IXb**	**Xa**	**Xb**
Academic Achievement	High School Grade Average	**-0.733****	-0.562	-0.466	-0.327	-0.272
	University Grade Average	**-1.066*****	**-1.529*****	**-1.499*****	**-1.487*****	**-1.656*****
Home Institution Characteristics	*Reference Category: University*					
	University of Applied Sciences	-/-	-0.808	-0.586	-0.348	-0.181
	Technical University	-/-	-0.013	-0.065	0.149	0.187
Field of Study	*Reference Category: Law, Economics and Social Sciences*					
	Engineering	-/-	-/-	-/-	-/-	0.432
	Mathematics, Informatics & Natural Sciences	-/-	-/-	-/-	-/-	0.483
	Linguistic and Cultural Sciences	-/-	-/-	-/-	-/-	-0.204
	Other Fields of Study	-/-	-/-	-/-	-/-	**2.607***
Language Proficiency	Language Skills	-/-	0.020	0.018	0.017	0.025
	TOEFL	-/-	0.003	0.169	0.109	0.041
Work Experience Part-Time	*Reference Category: No Part-Time Job*					
	at University	-/-	-0.463	0.445	0.445	0.686
	both at University & outside University	-/-	-0.125	-0.056	0.001	0.117
	Outside University	-/-	0.087	-0.188	0.154	0.270
Recipient of other Sponsorships	'BAföG'	-/-	0.176	0.158	0.067	0.179
	Any other Merit-Based Stipend	-/-	0.064	-/-	-/-	-/-
	'Studienstiftung des Deutschen Volkes'	-/-	-/-	**1.574***	**1.829***	**1.961***
	Applied for other Scholarships	-/-	**0.731***	**0.742***	**0.919****	**0.979****
Project-Specific Statements	Tuition Fees (in Thousand US)	-/-	-0.018	-0.019	-0.020	**-0.023***
	Private Guest Institution	-/-	0.696	0.724	0.890	0.902
	Guest Institution=Top10 University	-/-	0.609	0.485	0.624	0.597
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*					
	Canada	-/-	-0.083	-0.008	0.217	0.112
	US-Midwest	-/-	0.846	0.795	0.984	1.031
	US-Northeast	-/-	-0.893	-0.937	-0.837	-0.937
	US-South	-/-	-0.215	-0.224	-0.204	-0.340
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*					
	One Type of ECAs	-/-	**1.058***	**0.994***	**1.144****	**1.165****
	More Types of ECAs	-/-	0.858	0.723	0.835	**0.960***
Letter of Reference	Recommendation	-/-	-0.272	-0.262	-0.317	-0.327
	Recommending Person=Professor	-/-	0.506	0.482	0.541	0.430
	Good Relation to Recommending Person	-/-	-0.619	-0.678	-0.694	-0.794
Applicant Characteristics	Gender	-/-	-/-	-/-	0.377	0.416
	Glasses	-/-	-/-	-/-	0.439	0.305
	Born in Germany	-/-	-/-	-/-	0.234	0.305
	Age	-/-	-/-	-/-	**-0.313***	**-0.302***
	Parents=Academics	-/-	-/-	-/-	0.147	0.292
Control Variables	*Semester*	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Participant in Organized Study-Abroad Program	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Selection Year	-/-	*incl.*	*incl.*	*incl.*	*incl.*
Constant		**0.401*****	-0.788	-0.859	-1.843	-1.916
Observations		226	226	226	226	226
Pseudo R^2		0.068	0.169	0.181	0.200	0.221
*significant at * p <0.10 ** p <0.05 *** p <0.01*						

Table 7-7 Logistic Regression Coefficients Estimations VIII-Xb

Goodness of Fit and Model Comparison	Estimation Number[125] VIII	IXa	IXb	Xa	Xb
Cases correctly classified	61.95%	69.91%	71.24%	71.24%	**75.22%**
Log-Likelihood Full Model	-142.303	-126.964	-125.079	-122.169	**-118.900**
LR	20.846	51.524	55.294	61.114	**67.651**
(Prob>LR)	(0.000)	(0.004)	(0.002)	(0.002)	**(0.002)**
Mc Fadden's (Adjusted) R^2	0.068	0.169	0.181	0.200	**0.221**
	(**0.049**)	(-0.021)	(-0.009)	(-0.023)	(-0.027)
Cox-Snell R^2	0.088	0.204	0.217	0.237	**0.259**
Nagelkerke R^2	0.119	0.275	0.293	0.320	**0.349**
McKelvey and Zavoina's R^2	0.115	0.303	0.357	0.397	**0.444**
(Adjusted) Count R^2	0.619	0.699	0.712	0.712	**0.752**
	(0.065)	(0.261)	(0.293)	(0.293)	**(0.391)**
AIC	**1.286**	1.380	1.364	1.382	1.388
AIC*n	**290.606**	311.928	308.158	312.337	313.800

Table 7-8 Goodness-of-Fit Measures Estimations VIII-Xb

Taking a look at table 7-8, the reader can identify that all of the presented models explain stipend awarding decisions better than the intercept (or null) model, as all Likelihood-Ratio (LR) tests lead to the rejection of the null-hypothesis. However, in comparison to the GoF measures of Estimations I to VIIb, all of the GoF measures presented here are quite disappointing. Even in the best model (Estimation Xb), the pseudo R^2 values are as low as 0.221 and most of the adjusted McFadden's R^2 are even negative, indicating a poor model fit. Using Estimation Xb coefficients, one is able to correctly classify 75.22% of all cases, but the adjusted Count R^2 shows that a large part of this correct classification would have also been achieved by simple guessing. As none of the models VIII-Xb explains final selection decisions on a satisfactory level, no marginal effects will be discussed. Apparently, further empirical specifications are needed that better account for the peculiarities of final selection processes.

Due to the relatively small number of observations in final selection (n=254), most non-significant influences from Estimations VIII to XXb have been dropped and further final-selection-specific regressors have been included in Estimation XI-XIII. These include interview panel characteristics (Est. XI-XIII), situational factors such as interview time (Est. XII & XIII) as well as possible evaluator expectations measured in terms of pre-selection outcomes (Est. XIII).

[125] Note that Estimation IXa is not nested in Estimation Xb.

$$L_{XI_i}(Y_{Aw} = 1|Y_{Inv} = 1)$$
$$= \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i$$
$$+ \beta_3 LanguageSkills_i + \beta_4 SponsorshipStudienstiftung_i$$
$$+ \beta_5 Sponsorshipapplication_i + \beta_6 Tuition_i + \beta_7 PrivateInstitution$$
$$+ \beta_8 THETop10 + \beta_9 Number\ of\ ECA + \beta_{10} LOR + \beta_{11} Control + \beta_{12} Gender_i$$
$$+ \beta_{13} PlaceofBirth + \beta_{14} Age + \beta_{15} Parents_i + \beta_{16} PanelSize_i$$
$$+ \beta_{17} FractionFemaleEvaluators_i + \beta_{18} EvaluatorAge_i + \beta_{19} SelectionYear_i$$
$$+ \varepsilon_i$$

$$L_{XII_i}(Y_{Aw} = 1|Y_{Inv} = 1)$$
$$= \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i$$
$$+ \beta_3 LanguageSkills_i + \beta_4 SponsorshipStudienstiftung_i$$
$$+ \beta_5 Sponsorshipapplication_i + \beta_6 Tuition_i + \beta_7 PrivateInstitution$$
$$+ \beta_8 THETop10 + \beta_9 Number\ of\ ECA + \beta_{10} LOR + \beta_{11} Control + \beta_{12} Gender_i$$
$$+ \beta_{13} PlaceofBirth + \beta_{14} Age + \beta_{15} Parents_i + \beta_{16} PanelSize_i$$
$$+ \beta_{17} FractionFemaleEvaluators_i + \beta_{18} EvaluatorAge_i + \beta_{19} SelectionYear_i$$
$$+ \beta_{20} InterviewTime_i + \varepsilon_i$$

$$L_{XIII_i}(Y_{Aw} = 1|Y_{Inv} = 1)$$
$$= \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i$$
$$+ \beta_3 LanguageSkills_i + \beta_4 SponsorshipStudienstiftung_i$$
$$+ \beta_5 Sponsorshipapplication_i + \beta_6 Tuition_i + \beta_7 PrivateInstitution$$
$$+ \beta_8 THETop10 + \beta_9 Number\ of\ ECA + \beta_{10} LOR + \beta_{11} Control + \beta_{12} Gender_i$$
$$+ \beta_{13} PlaceofBirth + \beta_{14} Age + \beta_{15} Parents_i + \beta_{16} PanelSize_i$$
$$+ \beta_{17} FractionFemaleEvaluators_i + \beta_{18} EvaluatorAge_i + \beta_{19} SelectionYear_i$$
$$+ \beta_{20} InterviewTime_i + \beta_{21} PreselectionScore_i + \varepsilon_i$$

Logistic regression coefficients for all these models are displayed in table 7-9 and respective GoF measures are shown in table 7-10.

Logistic Regression Coefficients		Estimation Number		
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**XI**	**XII**	**XIII**
Academic Achievement	High School Grade Average	-0.509	-0.483	-0.140
	University Grade Average	**-1.624*** **	**-1.821*** **	**-1.817*** **
Language Proficiency	Language Skills	0.014	0.018	0.016
Recipient of other Sponsorships	'Studienstiftung des Deutschen Volkes'	**1.770** **	**2.045** **	**1.895** **
	Applied for other Scholarships	**0.786** **	**0.934** **	**0.840** **
Project-Specific Statements	Tuition Fees (in Thousand US)	**-0.021***	**-0.022***	-0.020
	Private Guest Institution	0.183	0.241	0.242
	Guest Institution=Top10 University	0.446	0.430	0.316
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*			
	One Type of ECAs	**0.969***	**0.940***	**1.046***
	More Types of ECAs	0.666	0.753	0.799
Letter of Reference	Recommendation	-0.310	-0.396	-0.551
	Recommending Person=Professor	0.480	0.308	0.257
	Good Relation to Recommending Person	-0.545	-0.489	-0.516
Applicant Characteristics	Gender	0.305	0.248	0.240
	Born in Germany	0.374	0.815	0.844
	Age	**-0.295***	-0.289	-0.304
	Parents=Academics	0.203	0.350	0.262
Evaluation Committee Characteristics	Size	-0.020	-0.200	-0.208
	Fraction of Female Evaluators	1.602	1.212	1.077
	Mainly Male Evaluators	0.410	0.516	0.341
	Average Evaluator Age	-0.045	-0.043	-0.052
	Dispersion Evaluator Age	-0.012	0.007	-0.016
Interview Framework	*Reference Category: Interview Time 10-10:59 a.m.*			
	Interview Time: 09-09:59 a.m.	-/-	0.051	0.032
	Interview Time: 11-11:59 a.m.	-/-	**-1.497** **	**-1.472** **
	Interview Time: 12-01:30 p.m.	-/-	-0.276	-0.323
	Interview Time: 02-02:59 p.m.	-/-	0.424	0.439
	Interview Time: 03-03:59 p.m.	-/-	-1.054	-1.170
	Interview Time: 04-04:59 p.m.	-/-	0.303	0.337
	Interview Time: 05-06:59 p.m.	-/-	-0.341	-0.343
Pre-selection Outcome	Pre-Selection Score	-/-	-/-	**0.126*** **
Year	Selection Year	**-0.697** **	**-0.851** **	**-0.892** **
Control Variables	*Semester*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	*incl.*	*incl.*	*incl.*
Constant		-1.559	-1.673	-1.356
Observations		226	226	226
Pseudo R^2		0.184	0.228	0.263
*significant at * p <0.10 ** p <0.05 *** p <0.01*				

Table 7-9 Logistic Regression Coefficients Estimations XI-XIII

Goodness of Fit and Model Comparison	Estimation Number		
	XI	XII	XIII
Cases correctly classified	74.34%	**74.78%**	73.89%
Log-Likelihood Full Model	-124.616	-117.842	**-112.624**
LR	56.220	69.768	**80.203**
(Prob>LR)	(0.000)	(0.000)	**(0.000)**
Mc Fadden's (Adjusted) R^2	0.184	0.228	**0.263**
	(0.014)	(0.012)	**(0.040)**
Cox-Snell R^2	0.220	0.266	**0.299**
Nagelkerke R^2	0.297	0.358	**0.403**
McKelvey and Zavoina's R^2	0.352	0.417	**0.482**
(Adjusted) Count R^2	0.743	**0.748**	0.739
	(0.370)	**(0.380)**	(0.359)
AIC	1.333	1.335	**1.298**
AIC*n	301.231	301.684	**293.249**

Table 7-10 Goodness-of-Fit Measures Estimations XI-XIII

Comparing the GoF measures presented in table 7-10 with those derived from Estimations VIII to Xb (table 7-8), one can see that Estimations XI to XIII better reflect the dynamics of the final selection process. Not only does the LR-test show that all of these models are significantly better than the intercept model. Most pseudo R^2 measures as well as the AIC values also point at a better model fit of Estimations XI to XIII. However, most pseudo R^2 measures are still comparably low, especially in comparison to table 7-5 and a large part of the correctly classified cases can be attributed to the baseline correct classification rate. Lower pseudo R^2 and GoF values in final selection than in pre-selection lead to the conclusion that final selection success cannot be modeled as precisely as pre-selection success (as expected in Hypothesis 9). Using the available applicant information, one is better able to predict pre-selection than final selection success. Among the available empirical specifications however, Estimation XIII provides the best fit and will subsequently be discussed in more detail.

Marginal Effects after Estimation XIII		dy/dx	Value of X
Academic Achievement	High School Grade Average	-.0287	1.55
	University Grade Average	**-.3717*** **	1.68
Language Proficiency	Language Skills	.0033	90.82
Recipient of other Sponsorships	'Studienstiftung des Deutschen Volkes'[i]	**.2298****	0
	Applied for other Scholarships[i]	**.1955****	1
Project-Specific Statements	Tuition Fees (in Thousand US)	-.0042	16.081
	Private Guest Institution[i]	.0468	0
	Guest Institution=Top10 University	.0601	0
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*		
	One Type of ECAs[i]	**.1630***	0
	More Types of ECAs[i]	.1854	1
Letter of Reference	Recommendation	-.1128	9.26
	Recommending Person=Professor[i]	.0553	1
	Good Relation to Recommending Person[i]	-.0932	1
Applicant Characteristics	Gender[i]	.0465	0
	Born in Germany[i]	.1965	1
	Age	**-.0621***	21.84
	Parents=Academics[i]	.0505	0
Evaluation Committee Characteristics	Size	-.0426	4.36
	Fraction of Female Evaluators	.2203	0.26
	Mainly Male Evaluators[i]	.0744	1
	Average Evaluator Age	-.0107	52.41
	Dispersion evaluator age	.0033	15.96
Interview Framework	*Reference Category: Interview Time 10-10:59 a.m.*		
	Interview Time: 09-09:59 a.m.[i]	-.0067	0
	Interview Time: 11-11:59 a.m.[i]	**-.3500****	0
	Interview Time: 12-01:30 p.m.[i]	-.0702	0
	Interview Time: 02-02:59 p.m.[i]	.0808	0
	Interview Time: 03-03:59 p.m.[i]	**-.2776***	0
	Interview Time: 04-04:59 p.m.[i]	.0637	0
	Interview Time: 05-06:59 p.m.[i]	-.0749	0
Pre-Selection Outcome	Pre-Selection Score	**.0258*****	84.98
Year	Selection Year[i]	**-.1453***	1
Control Variables	*Semester*	incl.	4.39
	Existing Contacts to Guest Institution	incl.	0
[i] *dy/dx is for discrete change of dummy variable from 0 to 1*			

Table 7-11 Marginal Effects for Standard Applicant after Estimation XIII

In final selection, only university – but not high school – grades do influence evaluator decisions significantly. Standard applicants have a 37.2 pps higher (conditional) probability of being awarded the scholarship than those with university grades averaging 2.68 (mean university grade (1.68) plus 1).[126] Recipients of the 'Studienstiftung des Deutschen Volkes' scholarship have an increased probability of also being awarded the present stipend (+0.23). Moreover, the mere fact that an applicant has also applied for other scholarships is associated with increased selection success probabilities (+0.20). In which ways these two variables are able to act as effective signals will be discussed in detail below (Chapter 7.3). Linking final selection success to the number of extracurricular activities applicants have indicated in their written applications, only the indication of exactly one type of ECAs is associated with a significantly higher final selection probability (+0.163).

While evaluation committee characteristics such as size, gender and age composition do not impact final selection outcomes, interview time partly does. Standard applicants being interviewed from 11 to 11:59 am (3 to 3:59 am) have a 35 (27.8) pps lower success probability than those interviewed from 10 to 10:59 am. A one-unit-increase in the assigned pre-selection score is associated with a 2.6 pps increase in final selection success probability. In contrast to pre-selection, the respective selection year also affects success probabilities. A standard applicant in 2008 has a 14.5 pps lower conditional probability of being awarded a stipend than the standard candidate in 2007.

[126] Note that a standard applicant in final selection possesses characteristics that differ from the ones of a standard applicant in pre-selection. For a comparison of these two standard applicants, see Appendices 4 and 5.

7.2.3 Combined Estimation

Another way of empirically dealing with both selection stages is one single, i.e. combined, estimation of pre- and final selection success. As it is possible to rank all three possible outcomes

1. rejection after pre-selection (assigned value of e.g. m=1),
2. invitation to interview, but rejection after interview (e.g. m=2) and
3. invitation to interview and stipend (e.g. m=3)

in ascending order, an ordered logistic model (OLM) of the following form is feasible:

$$Pr(y \leq m|x) = F(\tau_m - X\beta_i) \text{ for } m = 1 \text{ to } J - 1.$$

However, for ordered regression models – both ordered logit and ordered probit – the parallel regression assumption[127] needs to be satisfied. This assumption states that the βs are equal for each value of m, i.e. the probability curves only differ in being shifted to the right or left, but do not differ in their slope (Long and Freese 2006). This assumption implies that the influence of any individual variable is the same for each category of Y^*. As the previous analyses have shown, various independent variables do not have the same impact in pre- as in final selection. Testing the proportional odds assumption confirms the violation of this prerequisite and hence ordered logit (and probit) models should not be estimated in this context. Nevertheless, in order to be able to estimate both pre- and final selection success simultaneously, Estimation XIV was conducted using logistic regression again. This estimation includes all regressors from Estimation III, but additionally models the influence of the individual's predicted probability of being invited to an interview ($\hat{P}_{Inv_i}$: calculated from Estimation VIIa) on the final stipend awarding decision.

$$\begin{aligned} L_{XIV_i}(Y_{Aw} = 1) = {} & \beta_0 + \beta_1 AcademicAchievement1_i + \beta_2 AcademicAchievement2_i \\ & + \beta_3 HomeInstitutionChar_i + \beta_4 LanguageSkills_i + \beta_5 TOEFL_i \\ & + \beta_6 WorkExperience_i + \beta_7 SponsorshipBaföG_i \\ & + \beta_8 SponsorshipStudienstiftung_i + \beta_9 Sponsorshipapplication_i + \beta_{10} Tuition_i \\ & + \beta_{11} PrivateInstitution + \beta_{12} THETop10 + \beta_{13} GuestInstitutionRegion \\ & + \beta_{14} Number\ of\ ECA + \beta_{15} LOR + \beta_{16} Gender_i + \beta_{17} Glasses_i \\ & + \beta_{18} PlaceofBirth + \beta_{19} Age + \beta_{20} Parents_i + \beta_{21} \hat{P}_{Inv_i} + \beta_{22} Control + \varepsilon_i \end{aligned}$$

[127] In logistic regression models, this assumption is also called the proportional odds assumption.

Logistic Regression Coefficients		Estimation Number	
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**III**	**XIV**
Academic Achievement	High School Grade Average	**-0.908*** **	-0.410
	University Grade Average	**-1.592*** **	**-1.046** **
Home Institution Characteristics	*Reference Category: University*		
	University of Applied Sciences	-0.673	-0.281
	Technical University	-0.507	-0.283
Language Proficiency	Language Skills	0.029	0.012
	TOEFL	**0.859*** **	0.242
Work Experience Part-Time	*Reference Category: No Part-Time Job*		
	at University	0.186	0.251
	both at University & outside University	0.147	-0.140
	Outside University	-0.009	0.023
Recipient of other Sponsorships	'BAföG'	0.118	0.254
	'Studienstiftung des Deutschen Volkes'	**1.164***	0.597
	Applied for other Scholarships	0.407	**0.538***
Project-Specific Statements	Tuition Fees (in Thousand US)	-0.004	-0.014
	Private Guest Institution	0.317	0.635
	Guest Institution=Top10 University	0.014	0.258
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*		
	Canada	0.235	0.168
	US-Midwest	0.513	0.518
	US-Northeast	-0.285	-0.585
	US-South	0.146	-0.022
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*		
	One Type of ECAs	**1.004** **	0.647
	More Types of ECAs	**0.952** **	0.524
Letter of Reference	Recommendation	0.231	-0.024
	Recommending Person=Professor	**1.014*** **	**0.696***
	Good Relation to Recommending Person	0.698	0.074
Applicant Characteristics	Gender	0.072	0.133
	Glasses	0.308	0.221
	Born in Germany	0.849	0.552
	Age	**-0.272***	-0.161
	Parents=Academics	-0.163	0.163
Pre-selection Success	Individual Invitation Probability	-/-	**2.789** **
Control Variables	*Semester*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	*incl.*	*incl.*
	Participant in Organized Study-Abroad Program	*incl.*	*incl.*
	Selection Year	*incl.*	*incl.*
Constant		**-5.151*** **	**-5.574*** **
Observations		429	423
Pseudo R^2		0.306	0.320
*significant at * p <0.10 ** p <0.05 *** p <0.01*			

Table 7-12 Logistic Regression Coefficients Estimations III and XIV

Goodness of Fit and Model Comparison	Estimation Number	
	III	XIV
Cases correctly classified	77.86%	**79.91%**
Log-Likelihood Full Model	-184.844	**-179.061**
LR	163.106	**168.592**
(Prob>LR)	(0.000)	**(0.000)**
Mc Fadden's (Adjusted) R^2	0.306	**0.320**
	(0.179)	**(0.187)**
Cox-Snell R^2	0.316	**0.329**
Nagelkerke R^2	0.445	**0.462**
McKelvey and Zavoina's R^2	**0.542**	**0.542**
(Adjusted) Count R^2	0.786	**0.799**
	(0.313)	**(0.361)**
AIC	1.020	**1.012**
*AIC*n*	437.688	**428.121**

Table 7-13 Goodness-of-Fit Measures Estimations III and XIV

Indeed, Estimation XIV is superior to Estimation III in terms of almost all GoF measures (see table 7-13). As expected, pre-selection success in terms of the individual predicted probability of being invited to an interview is highly correlated with overall success. Most other significant influences found in Estimation III are no longer of importance as soon as pre-selection success is included in the model. Only university grades, having also applied for other scholarships and providing an LOR issued by a full professor still significantly influence selection success over and above pre-selection success. Hence, Estimation XIV presents the ultimate robustness check for all the aforementioned estimations and provides strong support for the decision to separately investigate each selection stage.

Consequently, both pre- and final selection success determinants will be discussed and compared in detail in the following section.

7.3 DISCUSSION

Analogous to the literature review, the discussion section is structured according to the respective influence factors and will first address the influence of diverse ability signals on scholarship awarding decisions (7.3.1), then discuss the impact of individual difference factors (7.3.2) as well as social (7.3.3) and situational or extraneous factors (7.3.4).

7.3.1 The Influence of Ability Signals

7.3.1.1 Educational Attainment

As anticipated, educational attainment and especially grades have a statistically significant impact on the awarding decision. This is not at all surprising as grades are designed to mirror a student's academic performance and/or potential. They represent a generally accepted standardized means of assessing performance. This facilitates signal interpretation. Evaluators are professors themselves and are familiar with the grading system and the grade average in their respective field of study (which is very similar to the field of study their applicants are in). The influence of high school grade averages might be explained analogously, as high school performance measures are the same all over Germany, and professors, i.e. the evaluators, can easily assess the student's intellectual ability and/or motivation to learn in school. All in all, grades reveal to be a very effective signal in scholarship selection processes.

As a consequence of the German grading system, the influence identified in the previously presented estimations is consistently negative: As higher values in the German grading system are associated with poorer grades, scholarship awarding probability decreases with increasing values of high school or university grades. Overall, a one-unit increase (i.e. from 1.76 to 2.76) in the high school grade average c.p. reduces the chance of this applicant to be awarded the scholarship to only 0.4 times $(e^{-0.908})$[128] the chance of an average applicant. For worse university grade averages, the impact is even more severe: Having achieved university grades averaging 'only' 2.89 instead of 1.89 (average applicant), reduces the odds of being awarded to be as low as 0.2. In line with expectations, grades have a stronger influence on pre-selection decisions than on final selection decisions. As outlined before, pre-selection decisions aim at selecting all theoretically suitable candidates on the one hand and eliminating inappropriate ones on the other hand. In final selection, however, all remaining candidates are expected to

[128] As mentioned before, logistic regression coefficients can be transformed into odds ratios by taking e^{β}.

have achieved a minimum grade and hence suitability level, and other factors such as personality, demeanor or eloquence influence an evaluator's decision more strongly during final selection. In the present case, both high school and university grades have an impact on pre-selection success, but only university grades – being more specific and relevant to a study-abroad purpose – affect final selection decisions. In comparison, a standard applicant's predicted probability of being successful changes as follows in a) pre- and b) final selection as a function of high school and university grades.

High School Grade Average	***Predicted Success Probability***		
	Pre-Selection	**Final Selection**	**Overall Success**
1.0 (excellent)	**0.91**	0.73	**0.36**
1.5	**0.83**	0.71	**0.26**
2.0 (good)	**0.71**	0.70	**0.18**
2.5	**0.54**	0.69	**0.13**
3.0 (satisfactory)	**0.36**	0.67	**0.08**
3.5	**0.20**	-	**0.06**
University Grade Average	***Predicted Success Probability***		
	Pre-Selection	**Final Selection**	**Overall Success**
1.0 (excellent)	**0.93**	**0.90**	**0.54**
1.5	**0.86**	**0.77**	**0.35**
2.0 (good)	**0.74**	**0.58**	**0.19**
2.5	**0.56**	**0.36**	**0.10**
3.0 (satisfactory)	**0.37**	**0.18**	**0.05**
3.5	**0.22**	**-**	**0.02**

Table 7-14 Predicted Success Probabilities dependent on Grades

For a standard applicant, the probability of a) being invited to an interview and b) awarded the scholarship changes as a function of high school and university grades as presented in the following conditional effect plots (figures 7-3 and 7-4).

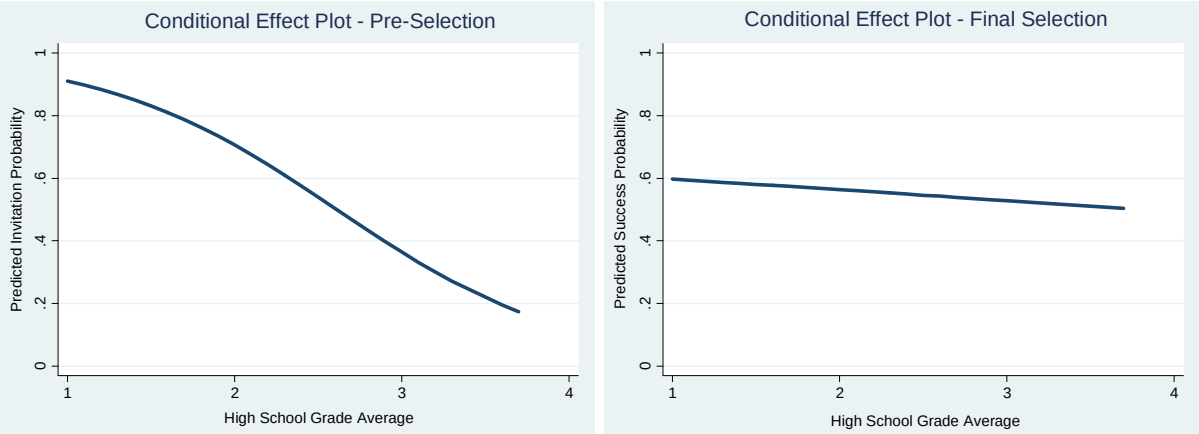

Figure 7-3 Conditional Effect Plots: High School Grades

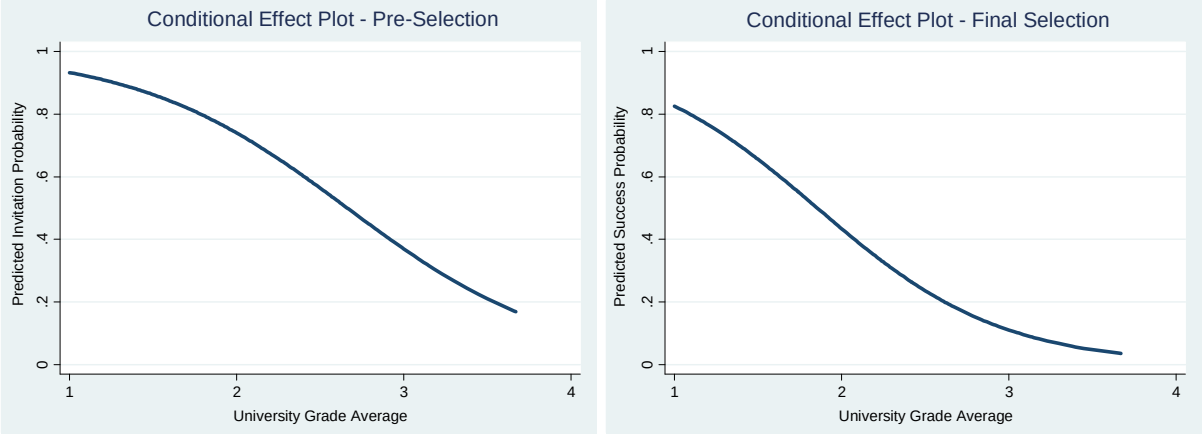

Figure 7-4 Conditional Effect Plots: University Grades

It is obvious from figure 7-3 that high school grades only affect pre-selection success significantly, whereas university grade averages (figure 7-4) are decisive in both, pre- and final selection. Overall, students with university grades worse than 2.0 (good) c.p. only have a probability of being awarded a stipend of 19% whereas students with outstanding grades will be awarded a stipend with a probability of 54%.[129] Hence, very good grades seem to be a necessary condition to be awarded a stipend.

In comparison to personnel selection research, grades consistently do play a more important role in stipend awarding decisions. This might be due to the fact that in education settings, grades are not only a productivity signal – as in recruitment – , but a direct productivity measure. All of the applicants wish to continue studying (both abroad and at home). Consequently, only the most 'productive' students are rewarded and productivity in this case is measured in terms of previous achievement, i.e. grades.

In derogation from the expectations, university grades are not only used as an initial screening method, but also affect final selection success. This might be caused by a relatively low variance in terms of other final applicant characteristics (such as ECAs, LORs or language skills) so that grades also serve as a means of comparison in final selection. In terms of other ability signals related to educational attainment, numerous other possible ability signals have been tested, but most have been demonstrated to not significantly influence either pre- or final selection outcomes. Among these are e.g. type

[129] Note that predicted probabilities as well as conditional effect plots are always calculated for a standard applicant whereas Odds Ratios are calculated from the aforementioned estimations, i.e. while fixing everything else at the mean. Consequently, results may at times seem inconsistent to the reader, but it needs to be noted that these differences result from the different computations.

and region of home institution, home institution reputation[130], desired degree (Bachelor vs. Diploma), length of study and number of institutions previously attended.[131] Differentiating between varying fields of study however reveals differing invitation probabilities as can be derived from Estimation VIIb. Both 'Engineering' and 'Linguistic and Cultural Science' students have significantly better chances (3.3 times and 2.6 times respectively) of being invited to an interview than 'Law, Economics or Social Sciences' students. This imbalance might be due to the fact that 'Engineering' is perceived as a difficult, yet desired field of study and pre-selection evaluators tend to invite more of these students to an interview. Additionally, as the overall grade average is worse in 'Engineering' (Wissenschaftsrat 2007), applicants with university grades averaging 1.89 (the mean university grade of all applicants) from this field of study are perceived to be better than applicants who have achieved the same grade average in other (less difficult) fields of studies. 'Linguistic and Cultural Science' students on the other hand might on paper seem especially suited for a study-abroad year as a result of their preparation in terms of content and language for instance. However, this effect only occurs in pre-selection. In final selection, students from all fields of study have comparable chances of being awarded a stipend (see Est. III).[132]

7.3.1.2 Extracurricular Activities

Extracurricular activities in general are positively associated with stipend awarding success. As almost all (84.3%) applicants state to be active in at least one area, further disaggregation is required. Consequently, the specific type of activity (political, clerical, social etc.) was measured as well as the number of different ECA types an applicant indicated to pursue. Additionally, the intensity of these activities was estimated from the available ECA information: The effort an applicant had with the indicated ECA was estimated based on a combination of (leadership) positions and assumed expenditure of time. This estimation led to the following categorization of ECA strength: high, medium and low strength (versus no ECA at all).

---

[130] Measured in terms of the CHE-ranking 2011 (research reputation as perceived by colleagues from related fields). Due to the inconsistent publication of CHE-results, this information was missing for approximately 50% of home institutions. Therefore, this regressor is not included in the aforementioned estimations and has only been used in several robustness checks. Results of all robustness checks are available on request.

[131] Furthermore, students from specific universities may be treated either more or less favorably than others. However, due to the limited sample size and the high number of different institutions (>60), further disaggregation was not possible and instead of testing for specific universities, institutions were only clustered and tested according to different characteristics, such as CHE reputation or type of institution.

[132] As the selection process in the investigated organization is structured according to fields of studies (of both applicants and evaluators), this is not surprising as the organization exogenously defines selection ratios which need to be complied with by each committee.

The number of ECA types indicated has a different effect in pre- and final selection. In pre-selection, the invitation chances of a candidate who is active in only one of the above mentioned ECA types are 2.7 times the chances of an applicant without any ECA mentioned in his or her CV. Mentioning more than one type of ECAs however multiplies this chance by 3.8. In final selection, however, the picture is reversed: Here, only those applicants who pursue exactly one (and no more) ECA have a significantly better chance of being awarded a stipend (2.8 times the chance of an applicant without any ECA). The differential effect of the number of ECAs pursued is also apparent when looking at predicted probabilities or conditional effect plots for both pre- and final selection.

Number of ECAs pursued	*Predicted Success Probability*		
	Pre-Selection	Final Selection	Overall Success
More than One Type of ECAs	0.77	0.71	0.22
Exactly One Type of ECAs	0.71	0.76	0.43
No ECA mentioned	0.48	0.52	0.10

Table 7-15 Predicted Success Probabilities dependent on Types of ECAs pursued

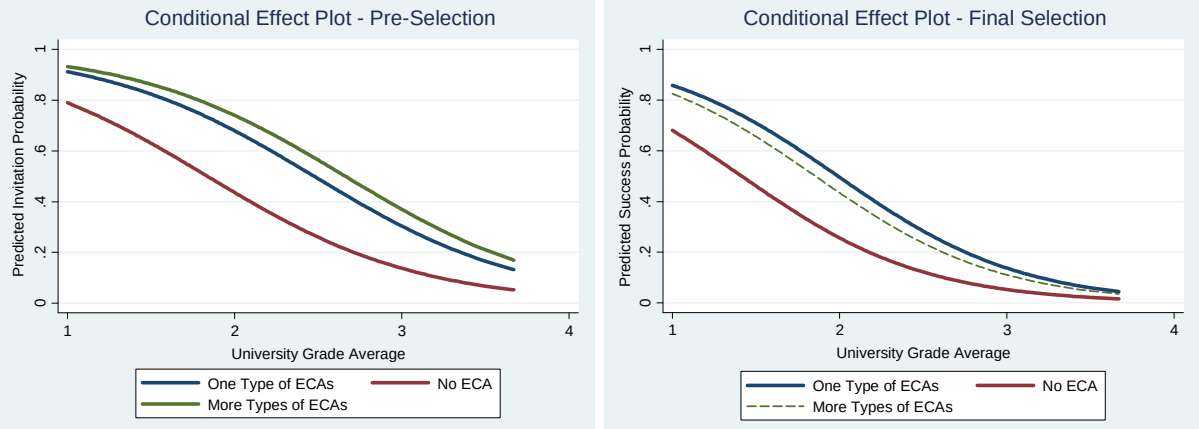

Figure 7-5 Conditional Effect Plots: Types of ECAs[133]

An explanation for the differential effect is straightforward: Whereas in paper applications, evaluators might be more impressed by a person who is (or at least indicates to be) active in various fields, final selection evaluators might favor the person who is only involved in one activity, but puts a lot of effort into this single activity. Applicants having indicated more than one type of ECAs in their CV might not be as active in reality as they pretend to be on paper. As it is not costly for applicants to indicate more activities on paper than they actually pursue, applicants are likely to polish their 'image' in an application. In final selection, however, it becomes more difficult (and costly) to pretend to be as active as indicated on paper. Candidates

[133] Note that insignificant effects will be illustrated with the help of a dashed line in all subsequent conditional effect plots.

pursuing one single activity are supposedly more committed to this activity and consequently more credible. Repeating Estimations III and VIIa with the variable 'Strength of ECAs' instead of 'Number of ECAs' confirms this impression: Whereas candidates with high, medium or low strength of ECAs all have significantly better (i.e. 4.3 times, 3.6 times and 3.5 times) chances of being invited to an interview, only those with high ECA strength are significantly more frequently awarded the scholarship in the end (7 times the chances of somebody without any ECA).

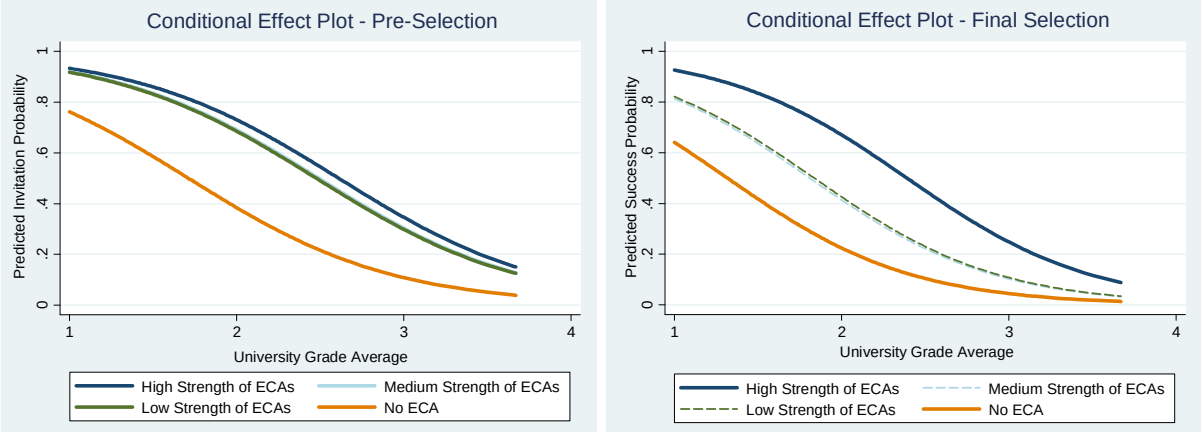

Figure 7-6 Conditonal Effect Plots: Strength of ECAs

Overall, applicants with exactly one type of ECA are slightly preferred during the stipend awarding process: Their chances of being successful are 2.7 times the chances of an applicant without ECAs, whereas applicants with more than one type 'only' have 2.6 times the chances of non-active applicants. In terms of strength of activity, the chances are as follows: highly active applicants have 5.8 times, medium active 3.2 and only slightly active candidates 3 times the chances of an inactive applicant. Taking a closer look at each type of activity independently reveals that only in pre-selection decisions the specific type of activity matters: Here, pursuing a clerical activity is associated with significantly lower (i.e. 0.5 times) chances of being invited to an interview, whereas applicants indicating to pursue extracurricular activities related to universities (student organizations and councils for instance) are invited more frequently (i.e. 2.2 times the chances of an applicant without this type of ECA). The rationale behind these two effects might be a slight similarity-attraction phenomenon as all evaluators are professors and consequently committed to university work. Hence, students supporting universities also during their 'free-time' might be favored due to a certain personal evaluator sympathy. A possible reason for treating applicants with clerical ECAs less favorably in pre-selection might be the independent nature of the investigated organization. Evaluators might be inclined to select applicants with clerical ECAs less frequently as these applicants might have access to other church-related stipends and are not as reliant on the current

scholarship as other candidates. However, neither in final selection nor in overall success, differential effects of any of the ECA types have been discovered. Thus, strength rather than type of activity seems to be rewarded by evaluators.

7.3.1.3 Letters of Reference

Standardized letters of reference (LORs) have some impact on success probabilities, but only in pre-selection decisions. In line with the expectations derived from personnel selection experience, LORs were in general very lenient. On the standardized scale from 0 (absolutely not suitable) to 10 (perfectly suitable candidate), recommenders assigned on average a value of 9.25 (range: 5 to 10). Due to the relatively low variance in recommendation values, it is not surprising that this standardized recommendation scale does not significantly affect overall success rates. In pre-selection however, a one-unit increase on this scale is associated with twice the chance of being invited to an interview. Consequently, variance in recommendation scores is reduced even more severely among final applicants and the recommendation itself does not influence final selection outcomes significantly. Presumably as a result of the low variance in recommendation scores, the recommending person per se as well as the indicated relationship between recommendee and recommender are associated with varying pre-selection outcomes. As depicted in figure 7-7, applicants who manage to receive a recommendation letter from a full professor (instead of an assistant professor or research assistant) are preferred in pre-selection. In the same vein, applicants whose recommenders indicate to know the applicant well are favored in pre-selection, too.

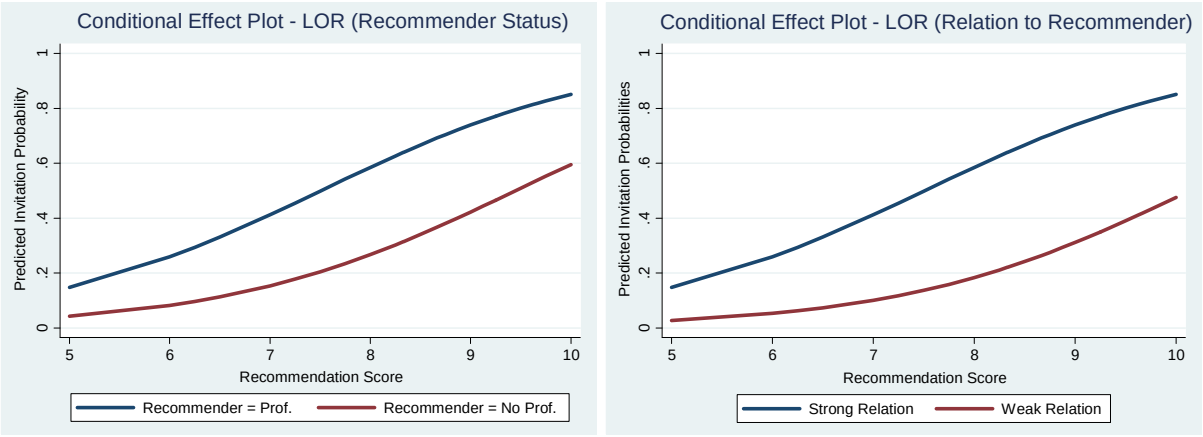

Figure 7-7 Conditional Effect Plots: LOR – Recommender Status and Relation to Recommender

Being able to get in touch with professors already at an early stage of their studies (2nd to 4th semester) seems to be interpreted as a credible ability signal. Evaluators supposedly know from their own experience as teaching professors how difficult it is for (undergraduate) students to contact full professors. Secondly, only very able students

will be recognized by professors in the classroom. Hence, persuading a professor to compose an LOR for him or her seems to be reserved for outstanding students. Consequently, this dummy variable is likely to measure some kind of unobservable heterogeneity. Applicants managing to have their recommendation letter written by a full professor seem to have at least some of the characteristics valued by evaluators. For whatever it exactly measures, the 'recommendation-written-by-a-professor' effect is so considerable that even in overall success rates applicants with letters of recommendation written by full professors have an increased stipend probability (2.8 times the chances of an applicant providing an LOR composed by assistant professors or research assistants).

As anticipated, LORs do not affect final selection success. As soon as the interview panel gets to know the candidate in person, LORs are no longer consulted. Furthermore, variance in recommendation scores, recommending person and relationship to recommendee is drastically reduced in final selection: scores only range between 8 and 10, 75.3% of all final selection applicants managed to get their LOR from a full professor and even 91.2% are well known by the recommender (according to the recommender's statements).

In comparison to personnel selection literature, stipend awarding evaluators tend to trust recommendation letters more than recruiters. A reason for this effect might be that professors (in related fields) do know each other better than recruiters in general know recommenders from other companies. Hence, evaluators in stipend selection processes rely more on the evaluation of their distinguished colleagues. Knowing the recommender well might then help 'reading between the lines' which is exactly the reason why LORs are usually not considered reliable in personnel settings.

7.3.1.4 Previous Awards

Evaluators in stipend awarding decisions might also look for heuristics in differentiating between appropriate and inappropriate candidates. One of these heuristics might be to consider previous awards or sponsorships the applicant has managed to receive. In the current context, previous sponsorships are directly inquired in the standardized application form so that this information is readily available to evaluators. Among these sponsorships are means-tested (e.g. BAföG) as well as merit-based scholarships. Overall, applicants indicating to have already been sponsored by another merit-based stipend[134] do not have significantly higher probabilities of receiving a stipend. However, being supported by one of the most renowned independent scholarship granting organizations, the 'Studienstiftung des Deutschen Volkes', is associated with extremely increased probabilities of being awarded the stipend at stake. This positive effect occurs both in pre- and in final selection, as can be derived from Estimations VIIa and XIII. Although stipend awardees from this institution are very likely to also have achieved excellent grades and strong ECAs, the previous-award effect occurs over and above the effect of these ability signals. Thus, recipients of the 'Studienstiftung des Deutschen Volkes' scholarship have 8.7 times the chance of being invited to an interview and 6.7 times the conditional chance of being subsequently awarded the stipend. Overall, these specific applicants have 3.2 times the chance of being awarded than applicants who have not been previously supported by the 'Studienstiftung des Deutschen Volkes'. Table 7-16 shows the 'Studienstiftung' effect in terms of predicted probabilities and figure 7-8 displays the distinct conditional effect plots for a) recipients of the 'Studienstiftung des Deutschen Volkes' and b) all other applicants as a function of university grades.

Supported by	*Predicted Success Probability*		
'Studienstiftung des Deutschen Volkes'?	Pre-Selection	Final Selection	Overall Success
Yes	0.97	0.94	0.47
No	0.77	0.71	0.22

Table 7-16 Predicted Success Probabilities dependent on Previous Award (Studienstiftung)

[134] E.g. from one of the organizations mentioned in Chapter 2.

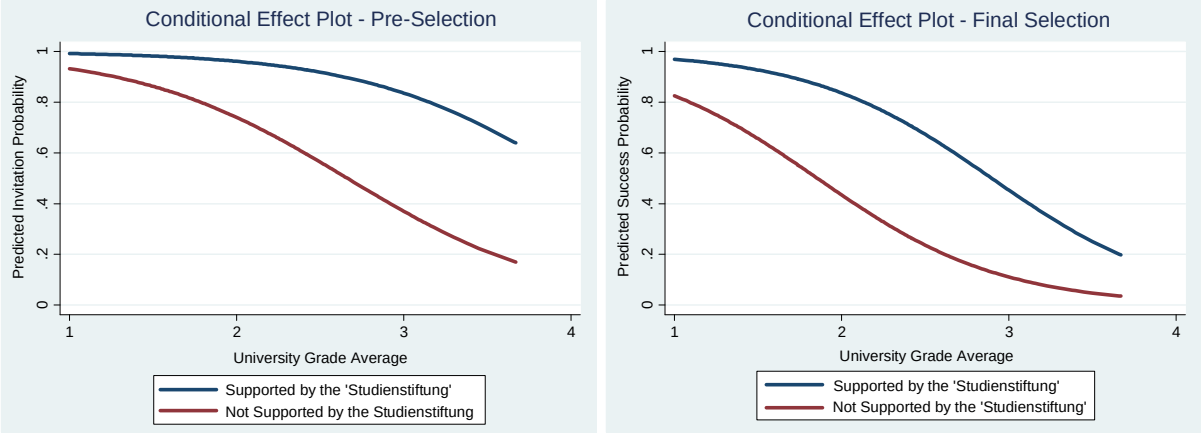

Figure 7-8 Conditional Effect Plots: 'Studienstiftung des Dt. Volkes'

Several explanations might explain this substantial effect. Either evaluators in the current context search for heuristics in order to facilitate selection. One obvious heuristic would be to simply select these applicants who have managed to already receive another stipend. Assuming that the other stipend-granting organization will have selected the right applicants during their selection process, evaluators in this context might rely on previous judgments or might simply be impressed by the previous achievement of this applicant. Both explanations are compatible with the theoretical notion of a 'Matthew-Effect' known from sociology: a cumulative advantage based on previous achievements (Merton 1968). On the other hand, the 'Studienstiftung' effect is likely to be the best proxy for a whole package of characteristics that all recipients of the 'Studienstiftung des Deutschen Volkes' possess. This combination of characteristics is likely to be desired by evaluators in the present context. In this case, the 'Studienstiftung' effect might be a very good measure of otherwise unobserved heterogeneity among applicants. Especially in final selection, the latter explanation is more likely. Applicants who have already managed to receive one of the most renowned stipends in Germany are very likely to be more self-confident and eloquent for instance – only two of the characteristics we are not able to observe during this research project. Hence, the 'Studienstiftung' effect is very likely to represent a combined measure of applicant quality.

Having been previously awarded a means-tested scholarship, i.e. BAföG, does not significantly influence stipend awarding probabilities. This is not surprising, as means-tested scholarships are usually not directly related to (previous) performance, but are assigned purely on the basis of an individual's financial means. Consequently, receiving a means-tested scholarship – in contrast to merit-based stipends – cannot function as a consistent ability signal.

7.3.1.5 Other Ability Signals

In addition to educational attainment, ECAs and LORs, previous work experience has been demonstrated to be one of the most important productivity signals in personnel selection. Assuming that past productivity is highly correlated with future productivity, recruiters infer the applicant's productivity potential from past achievements. In the current context, however, undergraduate students with an average age of 21 years are investigated, so that work experience is usually limited.[135] As a proxy for full-time work experience, both part-time work experience and the number and length of previously completed internships were tested. As usual, information was extracted from the applicants' CVs. The mere fact of having listed any part-time job at all in their CV as well as the total number of part-time jobs pursued until the time of application do neither affect pre- nor final selection outcomes.[136] Accordingly, evaluators do not seem to take into account the (financial) neediness when selecting appropriate candidates which is a considerable contrast to means-tested sponsorships. However, when differentiating between the type of part-time job employer, a significant positive effect for students having already worked part-time both at the university and for an external employer has been found in pre-selection: This applicant c.p. has 3.4 times the chance of being invited to an interview than an applicant who has not had any part-time job at all. Having worked either at university or for an external employer, does not change invitation probabilities significantly (see figure 7-9).

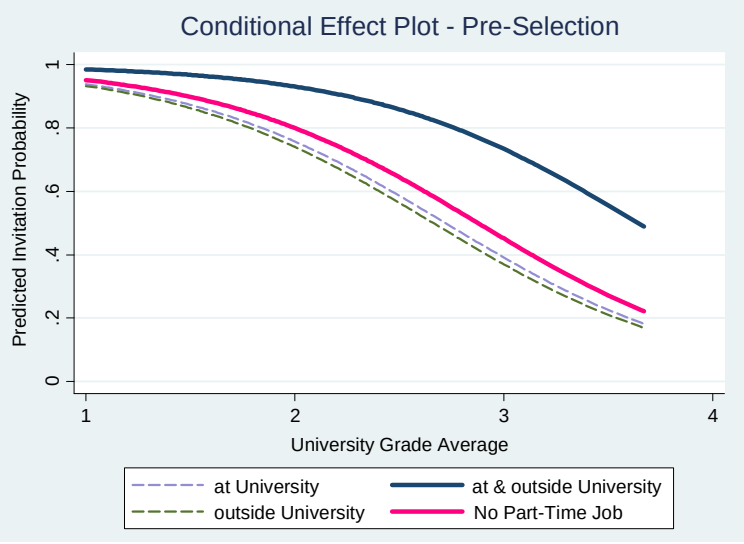

Figure 7-9 Conditional Effect Plot: Type of Part-Time Employer

[135] Very few students in the sample had indeed accomplished an apprenticeship prior to their studies – which can be understood as full-time work experience –, but when testing for the effect of apprenticeship (yes/no), no significant effects were found.

[136] Note that the measure of part-time jobs is a cumulative one and measures the total number of part-time jobs applicants have ever had in their lives. Whether or not several part-time jobs have been pursued simultaneously cannot be extracted from the available data.

In final and overall selection, however, this effect does not occur as can be derived from Estimations III and Xa. As the total number of part-time jobs pursued has no significant effect on selection outcomes, financial need per se does not serve as an explanation for the increased chance of an applicant with part-time occupations both at and outside the university. Instead, students who have already had diverse part-time jobs and still manage to achieve good grades might be perceived by evaluators as being more able to withstand stress and hence more suited for becoming an elite member in the future. As soon as evaluators however meet the 'paper' person face-to-face in final selection during the interview, other measures of resilience can be consulted and the part-time job proxy for the ability to handle stress is no longer important.

Another ability signal provided during this application process is foreign language proficiency. In the current context, language skills are measured in terms of a standardized scale from 0 to 100. This measure has been developed in retrospect using a combination of different proofs of language competence handed in by the applicants. As the current stipend is bound to a study-abroad period of at least nine months, language skills are expected to have a substantial influence on application success. However, only in pre-selection, the certified language skills have a significant effect on success: A one-unit increase on the language-skills scale is associated with a multiplication of invitation chances by 1.05. In final (and overall) selection, language proficiency as measured by a certificate or test does not impact success probabilities at all. Either written language certificates do not correlate substantially with real language skills (which might be tested during interviews) or variance in language skills is too low in final selection to have a substantial effect on awarding decisions. However, one effect which is related to language certificates is indeed apparent also in overall selection. Applicants handing in a TOEFL do have 2.4 times the chance of being awarded a scholarship than applicants handing in other language certificates. In pre-selection only, this effect is even more pronounced as can be derived from the predicted success probabilities of a standard applicant who either handed in a TOEFL or another certificate summarized in table 7-17.

Handed in TOEFL?	*Predicted Success Probability*		
	Pre-Selection	Final Selection	Overall Success
Yes	**0.77**	0.43	**0.22**
No	**0.44**	0.40	**0.11**

Table 7-17 Predicted Success Probabilities dependent on TOEFL

Again, several explanations are feasible for this effect. Either evaluators actually prefer applicants' handing in a standardized and objective test over applicants who provide only language certificates that have been issued by their own university. The rationale behind this reaction is that evaluators can more easily interpret the standardized results. On the other hand, applicants having already completed a TOEFL more than one year in advance of their stay abroad, might simply be perceived as more motivated and committed to the study-abroad project – characteristics that elite members should possess. Thirdly, a similar effect as discussed for all scholarship recipients of the 'Studienstiftung des Deutschen Volkes' in the applicant pool might be the actual driver of an increased invitation and awarding probability: Applicants who do neither spare the effort nor the cost of providing such a credible language-skill signal as the TOEFL might also have other characteristics that cannot be observed in the present context, but which positively affect selection success (unobserved heterogeneity). These applicants might for instance be more thorough in writing their motivation letter or describing their project. Without access to more applicant information, it cannot be concluded with certainty what drives the TOEFL effect, but it can be stated that applicants handing in such a test instead of another proof of language competence have increased chances of being awarded the stipend.

In addition to work experience and language skills, also facts and statements concerning the specific study-abroad project might act as ability signals. Evaluators might infer a student's ability and academic potential not only from past performance indicators, but also from a student's (study-abroad) intention and ambition. Therefore, it has also been tested whether or not the choice of the respective guest institution is likely to influence selection success. Again, due to the large number of different universities applicants plan to attend, guest institutions were clustered according to the following characteristics: guest institution region, reputation[137], funding[138], elite status[139] as well as the amount of tuition the student expects to pay for the entire year abroad. Among these, none of the project specifics significantly impacts overall success. In pre-selection however, applicants wanting to spend their year abroad at a public university are invited to an interview more frequently than those aiming at attending a private university (odds ratio: 0.4). The same applies to highly-ranked institutions: Indicating to plan studying at a university which is ranked among the top 10 of all North American institutions in the respective field, is associated with a decreased invitation probability

[137] Measured as the subject-specific THE-ranking 2010/11 (in terms of both points and rank). For more information, visit http://www.timeshighereducation.co.uk/.
[138] Privately or publicly funded.
[139] Ivy League institution versus all others.

(odds ratio: 0.5). The amount of expected tuition has a positive impact in pre-selection which is, however, not significant in Estimation VIIa. The slight positive tendency of increased tuition fees is likely to be caused by some outliers. Taking the natural logarithm of tuition fees instead or simply running the estimation without the top percentile leads to a non-significant impact in all of the above mentioned estimations. Applicants indicating their wish to study at an institution located in the Northeast of the United States are invited significantly more frequently than applicants planning to spend a study-abroad year at the West Coast. This effect persists although guest institution 'quality' in terms of reputation is controlled for.[140] Hence, the only plausible explanation for this differential effect is an evaluator's scepticism towards applicants wanting to spend nine months at the US-West coast. Serious study intentions are more likely to be expected from someone choosing a university in the Northeast than from someone planning to go to Hawaii or California for instance. The following conditional effect plots illustrate the differential pre-selection success rates dependent on guest institution region. It can indeed be derived from figure 7-10 that applicants planning to go to the US-West have the lowest selection probability.

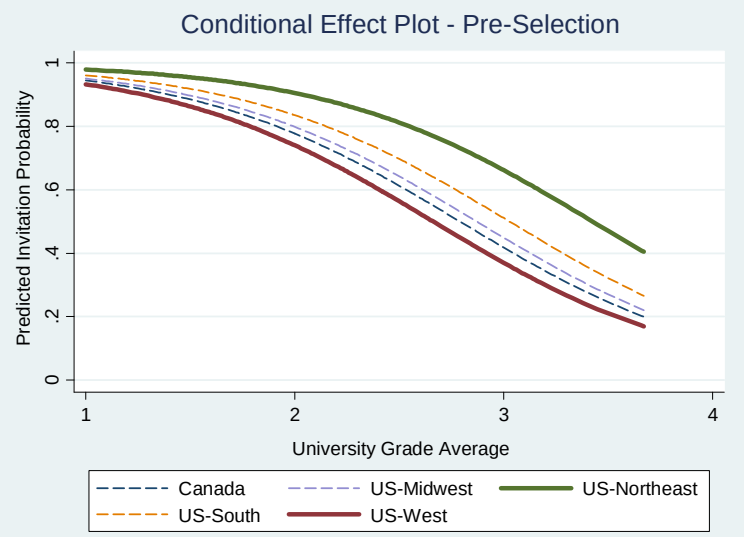

Figure 7-10 Conditional Effect Plot: Guest Institution Region

In final selection however, none of the above mentioned variables has a statistically significant impact on success rates. Hence, most project specifics rather seem to be taken into account during paper pre-selection, but do not influence awarding decisions in the end.

[140] As the most prestigious universities are traditionally located in the US-Northeast, one could assume that the region dummy instead measures prestige.

7.3.1.6 Interaction Effects: The Additive Effect of Ability Signals

As explained above, ability signals are not only likely to have a discrete main effect on selection success, but are also supposed to have a combined additive effect, i.e. these signals are supposed to be complements rather than substitutes. Hence, several interactions of ability signals that have been revealed in the preceding analysis to influence selection outcomes independently have been tested jointly.[141] As interaction effects in non-linear models vary as a function of all other independent variables, i.e. their value and significance is conditional on the values of other regressors, they cannot be interpreted globally as one would do in linear models (Ai and Norton 2003; Norton, Wang and Ai 2004). Consequently, marginal effects cannot be used, as they also vary according to the respective position and slope of the distribution in non-linear models. Hence, interaction effects derived from logit models have to be interpreted using graphs depicting the interaction effect for each observation, i.e. each combination of independent variables, independently. In other words, for each probability of being awarded the stipend (as predicted by all other independent variables) there is a specific interaction effect.[142] For the sake of brevity, only the graphs depicting at least some significant interaction effects will be presented subsequently.[143]

The interaction effect of number of ECAs pursued and university grade averages is non-significant. Replacing number of ECAs with ECA strength yields at least some significant positive interaction terms for applicants having an otherwise predicted probability around 0.6 to 0.8 and negative interaction effects for applicants with a very low probability of being awarded, as can be derived from figures 7-11 to 7-13.

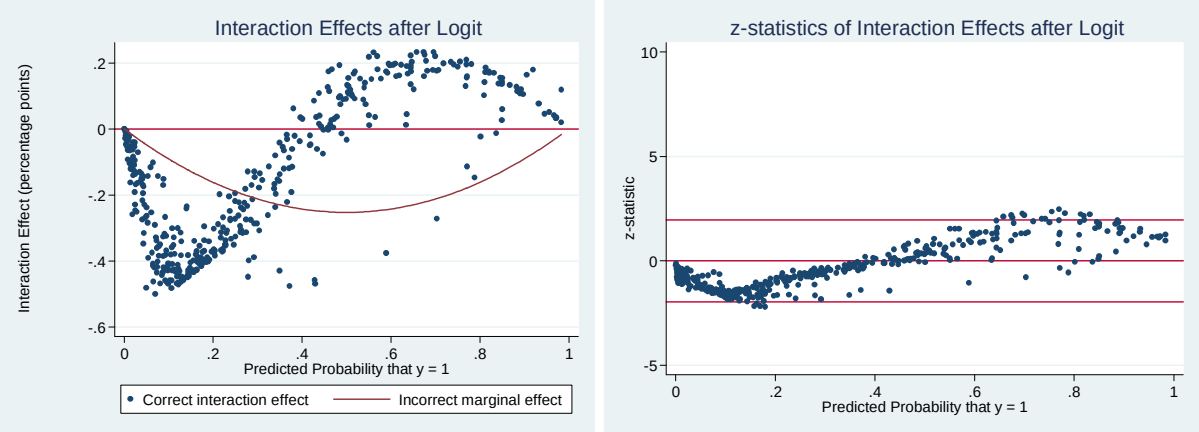

Figure 7-11 Interaction between University Grade Average and High Strength of ECAs

[141] For the sake of brevity, only results of interactions tested in Est. III, i.e. the overall success, will be presented. Results of interaction effects for other estimations are available on request.
[142] Graphs have been created using the *inteff* command suggested by Norton, Wang and Ai (2004).
[143] The results for non-significant interactions are however also available on request.

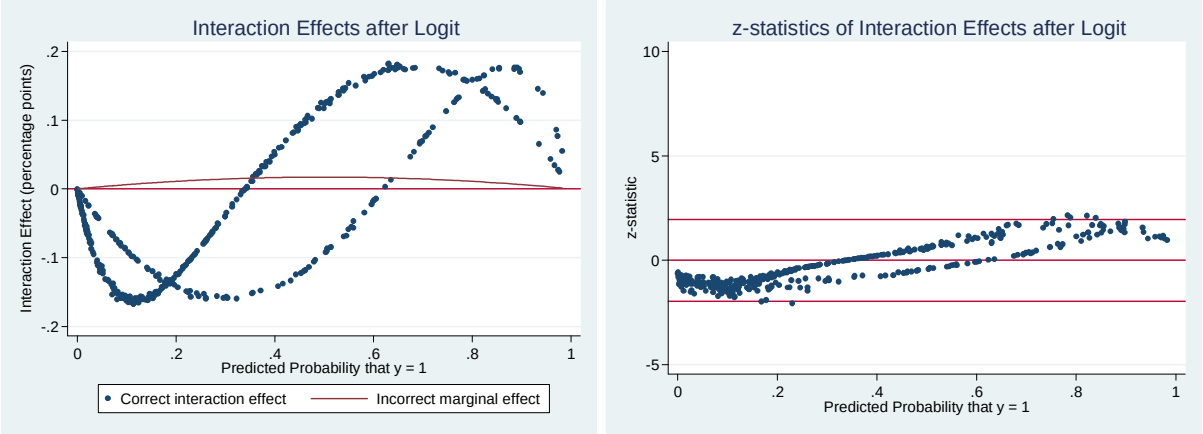

Figure 7-12 Interaction between University Grade Average and Medium Strength of ECAs

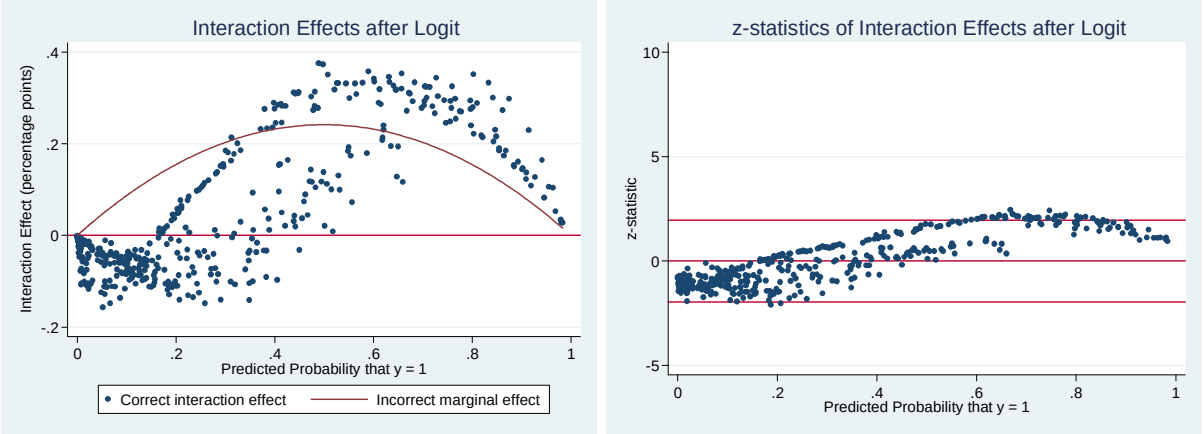

Figure 7-13 Interaction between University Grade Average and Low Strength of ECAs

When interacting the fact that someone is supported by the 'Studienstiftung des Deutschen Volkes' with university grade averages, the following result is obtained.

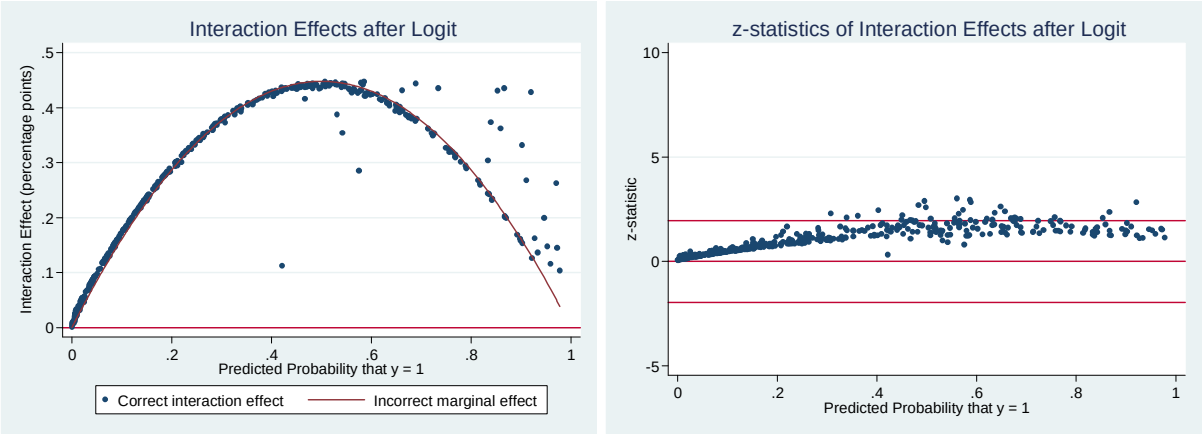

Figure 7-14 Interaction between University Grade Average and 'Studienstiftung'

Only for some of the applicants having a predicted probability ranging between 0.3 and 0.7, the interaction effect of Studienstiftung*University Grades reveals to be significantly positive. Recommendation interacted with university grades does not yield any significant interaction term.

Trying to find out whether the influence of university grades varies by type of institution, university grades were interacted with home institution type. No significant effect was found for technical university students, but for university of applied sciences students having an otherwise low probability of being selected ($0.1 < P < 0.4$) the interaction effect is significantly positive as figure 7-15 shows.

Figure 7-15 Interaction between University Grade Average and University of Applied Science

Examining the interaction between university grades and field of study, one might expect significant results, as university grades are traditionally better in some fields of study than in others. However, the only significant interaction which can be found in this context is the interaction between 'Linguistic/Cultural Sciences' and university grades. As depicted in figure 7-16, this effect is significantly positive for most of the applicants whose predicted probability is between 0.2 and 0.8.

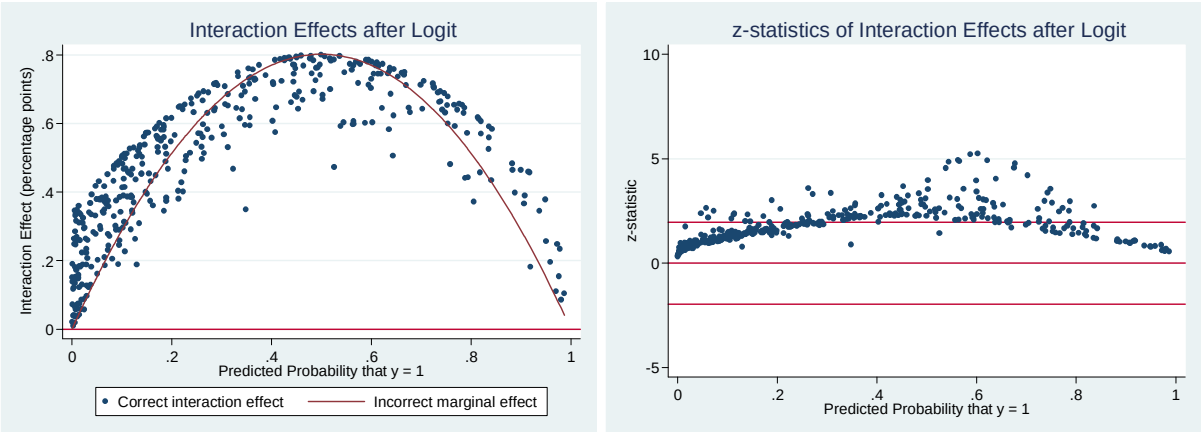

Figure 7-16 Interaction between University Grade Average and Field=Linguistic/Cultural Sciences

In summary and contrary to the expectations, most ability signals do not interact with each other. Each ability signal influences the success per se (main effect), but this influence is not reinforced or mitigated by changes in another ability signal (interaction effect). Thus, the complementary effect of ability signals cannot be supported.

7.3.2 The Influence of Individual Difference Factors

As anticipated in the conceptual model, individual difference factors should not affect rational decision making. Neither applicant characteristics that are unrelated to productivity – indices in Spence's (1973) terminology – nor rater characteristics should lead to different selection outcomes. However, in order to avoid omitted variable biases, several applicant and rater characteristics were included in the empirical estimation of (pre-)selection success (compare Hypothesis 5 and 6).

7.3.2.1 Applicant Characteristics

Applicant gender, ethnicity, age and physical attractiveness have been revealed to be the most important sources of discrimination in personnel selection.[144] Consequently, their influence was tested in the present empirical analysis. Applicant physical attractiveness could not be measured due to applicant data protection rights. Hence, only the influence of whether or not an applicant was wearing glasses on the picture attached to the CV could be tested empirically. Nevertheless, both pre- and final selection evaluators were exposed to photos and saw the candidate in person respectively. Accordingly, physical attractiveness might have an impact on awarding decisions, but it was not possible to measure this effect with the available data.

As anticipated, no applicant was discriminated due to gender. Neither in pre- nor in final selection, applicants of a specific gender were treated significantly more or less favorably, as can be derived from the very similar conditional effect plots in figure 7-17.

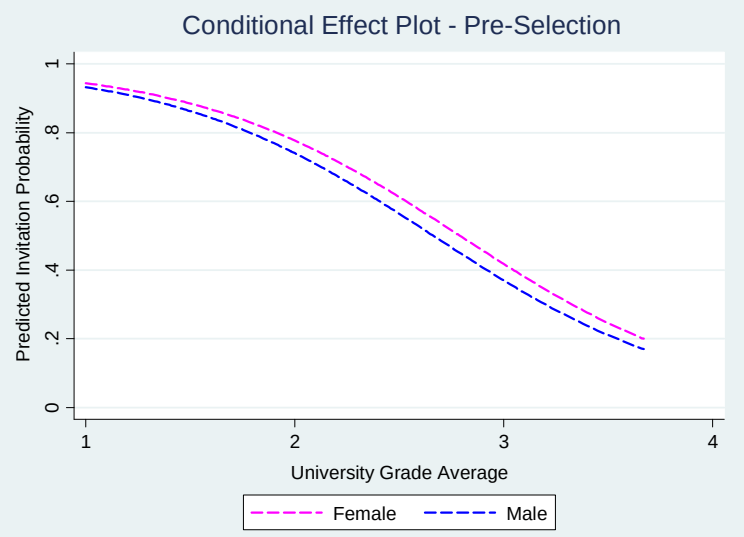

Figure 7-17 Conditional Effect Plot: Applicant Gender

The same applies to applicants with or without glasses. In line with the expectations, wearing glasses does not affect invitation or final selection probabilities.

[144] See Chapter 3.3.2.1 for a literature review.

Ethnicity could only be tested with the use of several proxies. Among these are the place of birth (in Germany or elsewhere), an applicant's last name (German-sounding, yes or no) as well as the candidate's citizenship (German, other or dual citizenship). Whereas the influence of the name itself and the citizenship revealed to be non-significant in all estimations, the country of birth did have a significant effect on invitation probabilities. Everything else equal, applicants born in Germany have 5.9 times the chance of 'surviving' pre-selection and being invited to an interview than a candidate who was born outside of Germany. The conditional effect plot as well as the predicted success probabilities vary substantially according to the country of birth as presented in figure 7-18 and table 7-18.

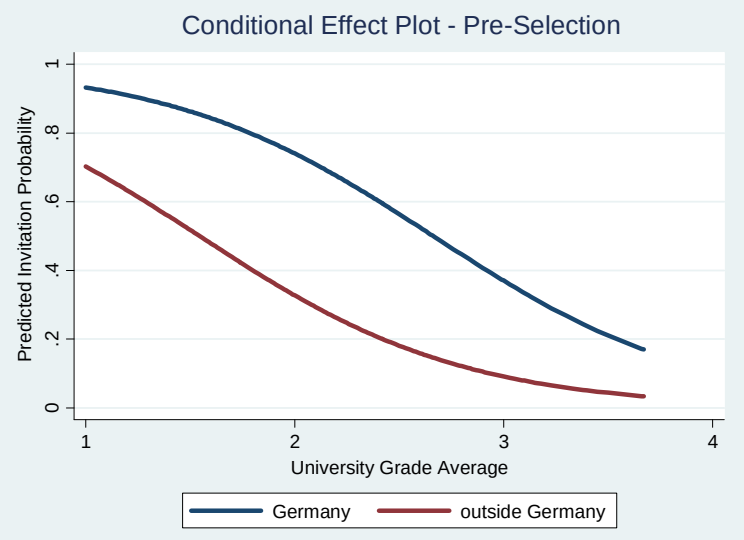

Figure 7-18 Conditional Effect Plot: Country of Birth

Born in Germany?	*Predicted Success Probability*		
	Pre-Selection	**Final Selection**	**Overall Success**
Yes	**0.77**	0.52	0.22
No	**0.37**	0.71	0.11

Table 7-18 Predicted Success Probabilities dependent on Country of Birth

The reader however needs to be reminded that only Germans or educational residents in Germany are entitled to apply for this specific program. Hence, the share of applicants who were not born in Germany is fairly low: Only 41 of all 504 applicants, i.e. 8.1%, were born in another country. Still, the effect is substantial for these 41 applicants. Standard applicants who were not born in Germany c.p. have a 40 pps lower probability of being invited to an interview. This difference in treatment does not occur in final selection and is not statistically significant in overall selection. Nevertheless, it needs to be analyzed

in more detail whether these applicants are in fact discriminated against (at least in pre-selection) or whether this group of applicants exhibits other, previously not considered attributes that reduce invitation probabilities. Therefore, a comparison of a series of other variables was conducted for the two groups of applicants who were a) born in Germany and b) not born in Germany. This comparison is presented below in table 7-19.

Variable	All Applicants	a)	b)	Difference (sign.)
Pre-Selection Score	75.42	75.78	71.37	**4.41 (**)**
High School Grade Average	1.80	1.80	1.85	-0.05 (+)
University Grade Average	1.94	1.93	2.00	-0.07 (+)
Semester	4.39	4.41	4.15	0.26 (+)
Home Institution in New Eastern State	5.20%	5.00%	7.30%	-2.3 Perc.pts. (+)
Recipient of 'BAföG'	22.60%	20.30%	48.80%	**-28.5 Perc.pts. (***)**
Previous Merit-Based Sponsorship	14.70%	14.70%	14.60%	0.1Perc.pts. (+)
Recipient 'Studienstiftung'	6.00%	6.00%	4.90%	1.1Perc.pts. (+)
Has Applied for Other Sponsorships	42.00%	41.90%	43.90%	-2Perc.pts. (+)
Private Guest Institution	27.60%	27.60%	26.80%	0.8Perc.pts. (+)
Tuition Fees at Guest Inst. (in Th. US$)	14.20	14.40	11.70	2.8 (+)
Participant in Study-Abroad-Program	27.80%	27.00%	36.60%	**-9.6 Perc.pts. (*)**
Guest Institution at US-Westcoast	21.23%	21.17%	21.95%	-0.8 Perc.pts. (+)
Female Applicants	48.60%	48.20%	53.70%	-5.5 Perc.pts. (+)
Professional Aim: Science/Research	19.80%	19.70%	22.00%	-2.3 Perc.pts. (+)
Age	21.90	21.90	21.50	0.4 (**)
Has Mentioned Parents in CV	33.50%	35.00%	17.10%	**17.9 Perc.pts. (***)**
Previous Stays Abroad	1.70	1.77	0.90	0.87 (***)
# of Completed Interships	1.50	1.52	1.20	0.32 (*)
# of Part-Time Jobs	1.92	1.92	1.95	-0.03 (+)
Total Length of Part-Time Jobs	28.60	28.40	31.20	-2.8 (+)
Extracurricular Activities (any)	84.30%	85.70%	68.30%	**17.4 Perc.pts. (***)**
# of LORs	1.27	1.27	1.29	-0.02 (+)
Recommendation in LOR	9.25	9.26	9.16	0.1 (+)
Recommender Knows Applicant Well	83.40%	83.30%	83.80%	-0.5 Perc.pts. (+)
Length of LOR	1.22	1.23	1.13	0.1 (+)
Recommender=Professor	68.10%	67.40%	75.60%	-8.2 Perc.pts. (+)
Certified Language Skills	88.60	88.60	88.80	-0.2 (+)
Has Handed in TOEFL	56.20%	57.50%	41.50%	**16 Perc.pts.(**)**
# of Additional Foreign Languages	1.86	1.82	2.32	**-0.5 (***)**
Differences significant on (*) 10%-Level; (**) 5%-Level; (***) 1%-Level; (+) not significant				

Table 7-19 Comparison of Characteristics according to Country of Birth

It is apparent that there are some significant differences between these two groups of applicants in terms of other variables that might affect (pre-) selection success. Applicants not born in Germany receive the most important means-tested scholarship in Germany (BAföG) significantly more often, indicating a lower socio-economic status of their parents. Hence, it is feasible that evaluators might have expected these students to receive a means-tested scholarship for their study-abroad project as well and did not invite them as a consequence of this assumption. Applicants not born in Germany also take part in organized study-abroad programs more frequently than applicants born in Germany. Furthermore, it is salient that these applicants less frequently mention their parents in the CV – something more than one third of applicants born in Germany do. In terms of ECAs, 'only' 68.3% of all applicants born in another country do indicate to pursue any of these activities. Moreover, they demonstrate their language skills less frequently with the help of a TOEFL than German-born applicants do. Not surprisingly, applicants born in another country do speak significantly more foreign languages than German-born applicants. This difference would rather explain an increased invitation probability for this group of applicants though.

Trying to answer the question whether these differences might drive the observable difference in treatment, it is obvious that most of them have already been included in the empirical estimations and are hence controlled for. Nonetheless, in order to empirically test whether actually these differences are driving the 'not-born-in-Germany' effect, several interactions were tested.[145] They all revealed to be either not significant or in the case of ECAs, only significant for applicants with a very high probability of being invited (based on other independent variables). Consequently, over and above the already mentioned differences, applicants who were not born in Germany are treated less favorably in pre-selection. It might however be that the applications from these candidates do contain something we were not able to measure in this research project, but which has an impact on pre-selection success. Thus, the dummy variable 'Born in Germany – yes or no' could simply capture otherwise unobserved heterogeneity (such as limited expressive powers both in speech and writing). As a result, we cannot for sure attribute the less favorable invitation probabilities of applicants who were not born in Germany to any kind of discrimination. On the other hand though, a convincing explanation has yet to be provided.

[145] Interactions tested: BAföG*Born in Germany, ECAs*Born in Germany, Study-Abroad Program*Born in Germany and TOEFL*born in Germany.

Age discrimination was also empirically tested: When holding the length of study (in terms of the number of semesters already completed) constant, younger applicants are consistently, i.e. both in pre- and in overall selection, favored. A student aged 22.8 instead of 21.8 years has 0.8 times the chance of being awarded the scholarship. This differential treatment is rooted in pre-selection, as the 'older' applicant has only 0.7 times the chance of being invited to an interview.[146] The decreasing invitation and overall selection probabilities as a function of applicant age are illustrated in figure 7-19.

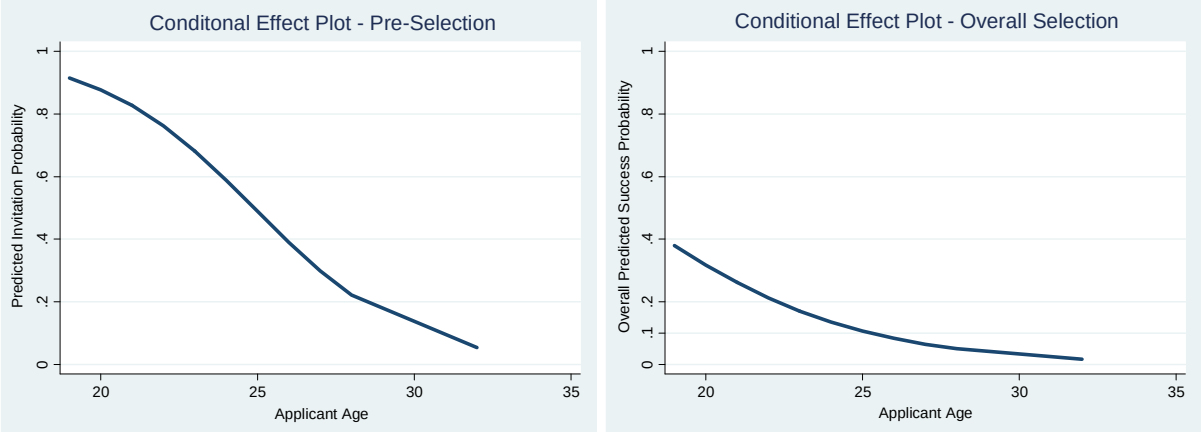

Figure 7-19 Conditional Effect Plots: Applicant Age

In the final selection however, applicant age does not significantly impact selection success. In my opinion, the preferred invitation of young(er) applicants can nonetheless not be understood as discriminatory practice in the same way it occurs (and has been empirically demonstrated) in personnel selection. As a consequence of their increased human capital in terms of work experience older job applicants are more suitable than young ones in most occupations. Younger students however having achieved the same as (slightly) older ones can be perceived as being more motivated and as possessing a high academic and non-academic potential. Thus, favoring younger applicants is likely to be based on other factors such as motivation and determination which are simply reflected in age and should not be labeled age discrimination per se.

[146] It has also been tested whether applicant age has a quadratic influence on pre-selection/ overall selection success. However, including Age^2 in the estimation resulted in a non-significant coefficient. It is indeed best for an applicant to be as young as possible (linear influence only).

The last applicant characteristic whose influence on selection success needs to be tested, is an applicant's socio-economic status. The goal here is to find out whether or not applicants coming from upper (middle) class families are preferred by evaluators.[147] Unfortunatly, information on an applicant's background was not consistently available. Only when applicants decided to mention their parents voluntarily in their CV,[148] information on the (presumable) socio-economic background could be gathered. As only one third of all applicants did mention their parents, it cannot be assumed that the variable 'Parents mentioned as Academics' precisely captures all applicants having an academic background, i.e. coming from families where at least one parent has graduated from a higher education institution. Nevertheless, this dummy was tested as a proxy for the academic-background effect. As can be read from table 7-20, mentioning academic parents in the CV has a significantly negative effect on pre-selection success, but not on final or overall success.

Mentioned Parents as Academics in CV?	*Predicted Success Probability*		
	Pre-Selection	Final Selection	Overall Success
Yes	**0.59**	0.71	0.19
No	**0.77**	0.76	0.22

Table 7-20 Predicted Success Probabilities dependent on Academic Background

Indeed, applicants who reveal themselves as coming from academic families have significantly lower chances of being invited to an interview. Consequently, evaluators might be prompted to favor an applicant from a non-academic background over an applicant whose background is known to be academic whenever these two applicants are otherwise equally suited (c.p.). The rationale behind this favoritism might be the lower anticipated socio-economic status of these applicants. Alternatively, evaluators could generally treat applicants less favorably who mention their parents in their CV as these might be perceived as immature and parent-focused. This impression is confirmed as replacing the variable 'Parents mentioned as Academic in CV' with the more general one 'Parents mentioned in CV' leads to similar negative results.[149] Whichever explanation is preferred by the reader, the assumption that applicants from academic backgrounds are treated more favorably by stipend granting evaluators has to be explicitly rejected following the empirical analysis. Thus, we can safely conclude that the results of Middendorff, Isserstedt and Kandulla (2009) were mainly driven by self-selection.

[147] This assumption was based on the findings of Middendorff, Isserstedt and Kandulla (2009) presented in Chapter 3.
[148] A practice that used to be quite common in German applications some decades ago.
[149] However, these two effects cannot be distinguished clearly, as more than 70% of all applicants mentioning their parents in the CV do have an academic background.

7.3.2.2 Rater Characteristics

Analogous to applicant characteristics, rater characteristics are not expected to influence rational decision processes. However, the literature review showed that e.g. rater gender may eventually have an effect on evaluations and sometimes even decision outcomes. Hence, rater characteristics were also included in the aforementioned estimations. In pre-selection situations, only one single evaluator decided on a paper application and hence his or her gender and age could be exactly measured. In line with previous empirical findings, female pre-selection evaluators revealed to be more lenient with the applicants they were assigned (table 7-21).

Evaluator Gender (Pre-Selection)	*Predicted Pre-selection Success Probability*
Female	**0.93**
Male	**0.77**

Table 7-21 Predicted Success Probability dependent on Evaluator Gender

Consequently, the six female pre-selection evaluators did not abide by the agreement of inviting only 50% of all applicants, but indeed invited 60% of all candidates they were asked to evaluate. This effect is global, i.e. unconditional on the applicant's gender.[150] In line with previous research findings in personnel selection, female evaluators tend to be not as rigorous with applicants as male evaluators are and want to give 'borderline' candidates a chance to present themselves in front of a committee. In final selection however, selection committees with a higher share of females did not award significantly more stipends than panels with less or even no female evaluators.[151]

In terms of pre-selection evaluator age, another significant effect was found. Older evaluators c.p. tend to be slightly more lenient with applicants as the conditional effect plot for a standard applicant as a function of evaluator age in figure 7-20 shows.[152] In final selection however, neither the average evaluator age nor the age disparity in a given panel leads to significantly different selection outcomes.

[150] An interaction between pre-selection evaluator and applicant gender was also tested, but will be discussed in section 7.3.3 (Social Factors).

[151] Again, interactions between share of female evaluators and applicant gender were included in the above mentioned estimations and will be discussed in section 7.3.3.

[152] Again, a quadratic influence of evaluator age was additionally tested, but did not have any influence on pre-selection success. The influence of (Evaluator Age)^2 was not significant.

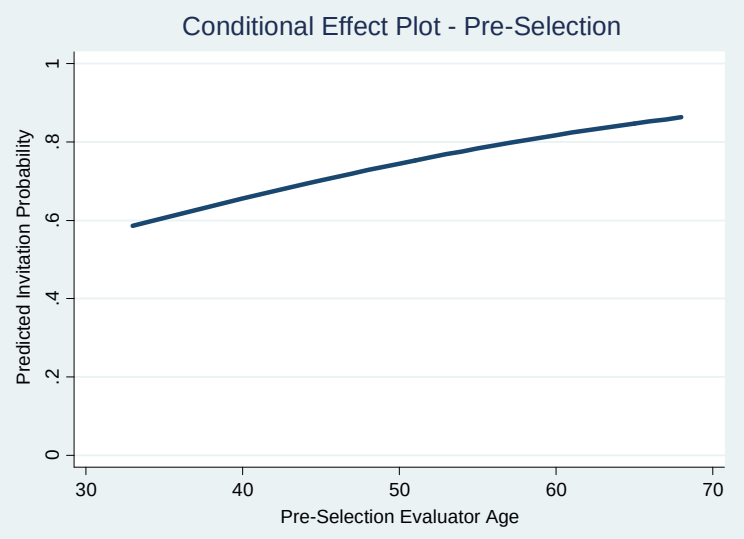

Figure 7-20 Conditional Effect Plot: Evaluator Age

7.3.3 The Influence of Social Factors

From the list of social factors addressed in section 3.3.3, only (demographic) applicant-rater similarity could be empirically tested with the available data. Although applicant fit and impression management tactics might have had a substantial influence on final evaluator decisions, they were simply not measured during interviews and could hence not be gathered in retrospect. Applicant-rater similarity could not be captured in terms of attitudinal similarity – as the original similarity-attraction paradigm hypothesizes – but had to be measured in different ways: gender, regional and institutional similarity. Gender similarity was modeled as an interaction term between applicant and rater gender as well as applicant gender and share of female evaluators in final selection panels, respectively. Regional similarity was modeled (for pre-selection only) as a dummy variable that equals 1 whenever the pre-selection evaluator teaches in the same federal state the applicant's home institution is located in. Institutional similarity refers to the type of higher education institution[153] the evaluator teaches at and the applicant attends, respectively. This type of similarity was also modeled as a dummy variable which equals one whenever the evaluator's and the applicant's institution were of the same type. Finally, a measure of overall applicant and evaluator similarity was introduced: As all evaluators are professors, it might be feasible that they perceive an applicant as more similar to them whenever he or she indicates to have a professional aim related to research or science. Such an aim has been indicated by 14.3% of all applicants.[154]

[153] University versus technical university versus university of applied sciences.

[154] The reader might be tempted to think about more similarity parameters, such as field of study and field of research, respectively. However, this similarity is given for all applicants, as the

Neither regional nor institutional similarity affected pre-selection outcomes significantly. Neither did the fact that an applicant stated to have a professional aim in research. Including the interaction effect between applicant and pre-selection evaluator gender into Estimation VIIa, yields the interaction effects depicted in figure 7-21. The z-statistic on the right hand side of figure 7-21 however shows that this interaction effect is not significant at any position in the distribution.[155] Hence, pre-selection evaluators do neither prefer candidates of their own gender over candidates of the opposite gender nor do they treat them less favorably.

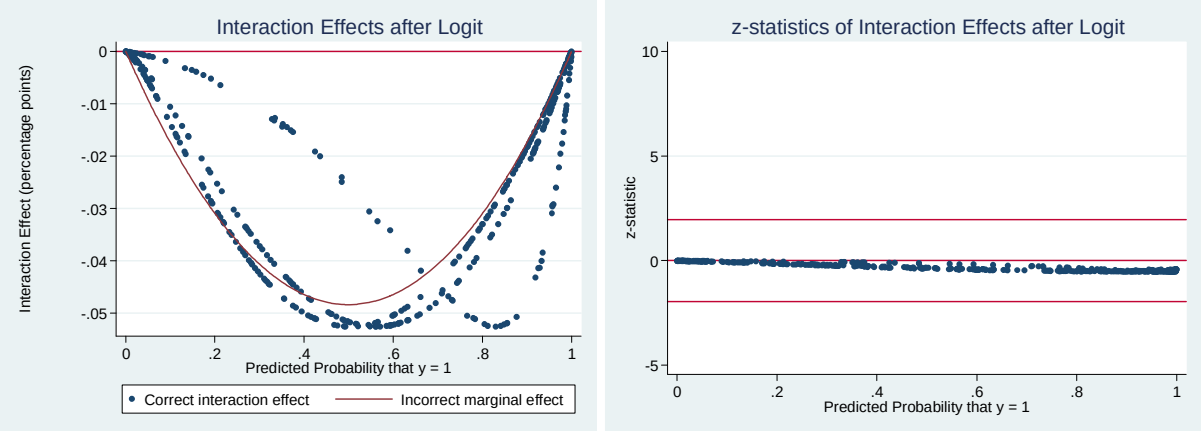

Figure 7-21 Interaction between Applicant and Evaluator Gender

Analogous to this pre-selection interaction effect, the interaction between the share of female evaluators and applicant gender in final selection was tested. The results of this interaction effect as well as the z-statistic are displayed in figure 7-22.

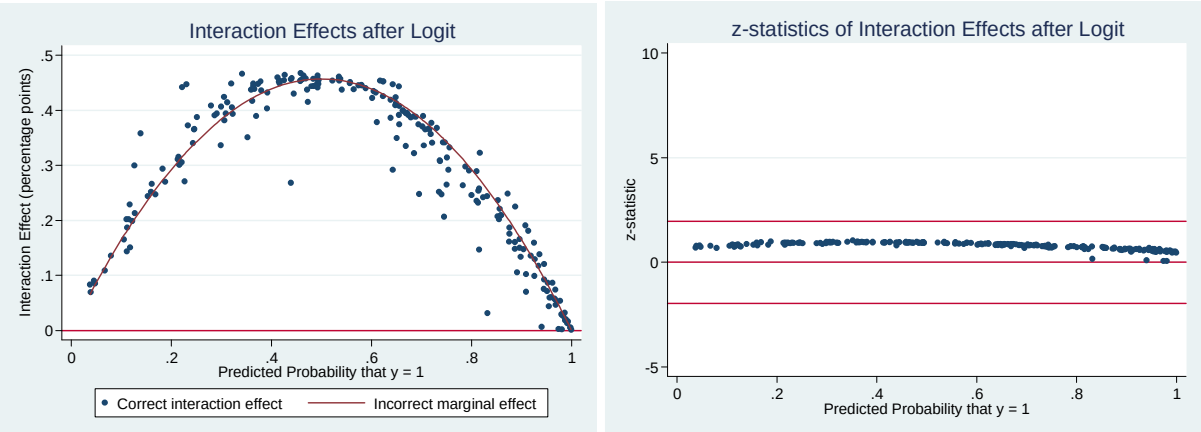

Figure 7-22 Interaction between Applicant Gender and Fraction of Female Panel Members

scholarship granting organization assigns applications to specific evaluators on the basis of the particular subject. Hence, all evaluators assess applications from students who are in the same field they work in.

[155] Both graphs have been created using the *inteff* command suggested by Norton, Wang and Ai (2004). This procedure is necessary as in non-linear models, the interaction effect varies according to the respective values of all other independent variables and can thus not be interpreted globally by simply looking at the marginal effect and its z-statistic created by including the interaction term in the estimation.

As can be derived from the z-statistic on the right hand side of figure 7-22, this interaction effect is not significant either. Hence, committees with higher shares of female evaluators do not treat female applicants more or less favorably than male applicants (and vice versa). Similarity-attraction in terms of gender was therefore detected neither in pre- nor in final selection.

All in all, evaluators were not susceptive to the similarity-attraction phenomenon at least as measured in the current context and therefore did not favor applicants who appeared similar to them in terms of gender, regional, institutional and biographical characteristics.

7.3.4 The Influence of Situational and Extraneous Factors

Similar to individual difference and social factors, situational or extraneous factors are not expected to influence rational decision processes. Among the available (and testable) situational factors are interview panel size and composition[156] as well as contrast effects[157] and the interview time for final selection. Furthermore, the provision of pre-selection scores can be understood as a situational factor as well, as it influences interviewer expectations. In pre-selection, the only extraneous factors that could be captured were evaluator gender and age – which have already been discussed in section 7.3.2.2 – and the number of applications which were allocated to this specific evaluator. The number of direct competitors did not affect pre-selection success significantly though.

As can be derived from Estimation XIII, panel size, i.e. the number of evaluators in the specific committee, did not influence final selection outcomes. Contrast effects did not occur either. However, the respective time of the day an interview is scheduled indeed affects selection success. An applicant being interviewed between 11 and 11:59 am, c.p. has only 0.23 times the chance of somebody being interviewed between 10 and 10:59 am (reference category). The differential success rates dependent on interview time are depicted in figure 7-23.

[156] Interview panel composition in terms of gender has already been discussed in sections 7.3.2.2 and 7.3.3.

[157] Contrast effects describe the influence of the immediately preceding candidate's performance on the success probability of the currently interviewed applicant. For more details on previous contrast effect research, see Chapter 3.3.4.

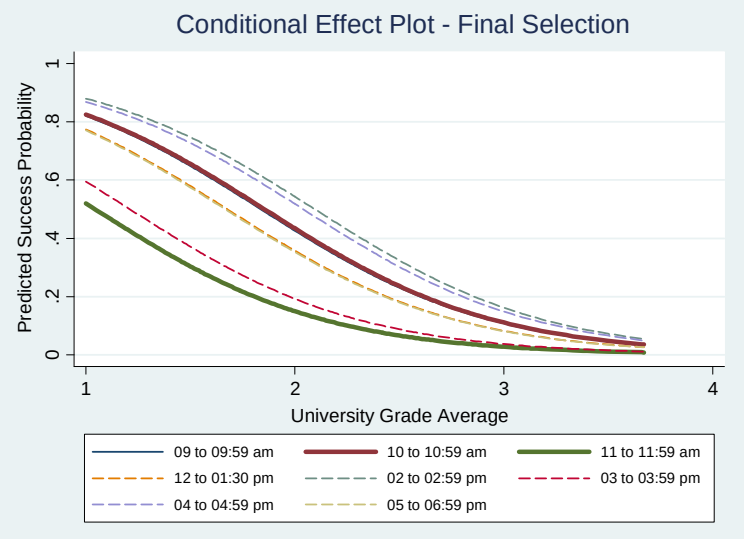

Figure 7-23 Conditional Effect Plot: Interview Time

In order to eliminate the possibility of a differential circadian distribution of applicant quality, several objective ability signals were compared for applicants being interviewed at different times of the day, but no significant differences in any of these quality dimensions were found. Hence, it indeed seems to be the extraneous influence of interview time[158] that decreases selection probabilities. Analogous to the findings by Danziger, Levav and Avnaim-Pesso (2011), this effect might be a symptom of evaluator fatigue and/or hunger after a long, uninterrupted duration of the meeting.

Although most of the factors associated with pre-selection success did not reveal to be significant in final selection, the pre-selection score itself does have a significant positive effect on final selection success as figure 7-24 shows.

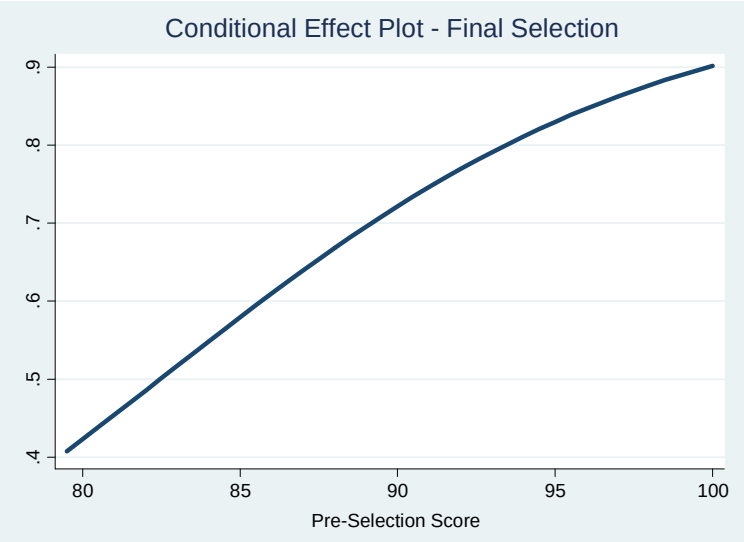

Figure 7-24 Conditional Effect Plot: Pre-Selection Score

[158] Or some other hitherto unobservable characteristics of applicants being invited at this time.

Accordingly, either applicants with a high pre-selection score do also possess characteristics that are related to final selection success, but could not be measured with the available data, or final selection evaluators do build expectations on the basis of pre-selection scores and consequently use this metric as decision support.

7.4 SUPPLEMENT - CONFIGURATIONS TO SUCCESS: BOOLEAN LOGIT REGRESSION

As anticipated in the conceptual model and revealed in the previous sections, not one single applicant characteristic is likely to be decisive for success or failure in a scholarship selection process. Several ability signals have been demonstrated to affect selection outcomes significantly, e.g. high school and university grades, recommendation and ECAs. However, most interactions between these different kinds of signals have not yielded significant results and reinforcing effects were not found. One reason for this might be the model specification. Simple logit regression is not able to model interaction effects that are neither linear nor additive in nature. Nevertheless, it is very likely that scholarship applicants need to possess a certain set of characteristics to be successful. For instance, it might be necessary, but not sufficient to have excellent grades in university. A high intensity of ECAs might be advantageous for candidates with particularly good grades, but this signal might per se not be strong enough for applicants with rather poor grades to differentiate from the rest of the applicant pool. Consequently, each single characteristic does not determine selection success, but certain configurations of several attributes and characteristics may be likely to lead to a positive outcome. This kind of causal complexity however needs to be examined using research methods other than simple logistic regression.

One approach to measuring the impact of different causal mechanisms on a given outcome is to apply qualitative comparative analysis (QCA). Using configurational comparative methods such as crisp set or fuzzy set QCA, the researcher is able to explore causal substitutability and can discover multiple paths to a given outcome (Ragin 1987). However, this approach is rather case-study oriented and qualitative in nature and hence better suited for a low number of observations (Buche and Carstensen 2009). On the other hand, conventional statistical, i.e. variable oriented, methods such as logistic regression are only able to identify the net effect of one variable (and some simple interactions of these variables) while holding everything else constant, i.e. in isolation, but provide robust results of the probable validity of postulated hypotheses. Whenever enough cases and enough variation within these observations exist, such correlational analysis can be conducted (Hellström 2011; Porcu, Puggioni and Tedesco 2007).

A method that combines the advantages of both statistical and configurational analyses is Boolean logit as suggested by Braumoeller (2003). It is a quantitative, i.e. variable oriented, method that is "designed to evaluate conditional or asymmetric causal claims" (Hellström 2011, 73) by the use of Boolean logic. In this analysis, the impact of various 'causal paths', i.e. configurations of variables, which together make the dependent variable occur, can be modeled with the help of the Boolean operators *and* and *or*. The impact of each causal path is in turn determined by some vector of independent variables. Therefore, it can be measured how "multiple causes interact with one another […], and the manner in which they interact is described by the logical operators 'and' and 'or' " (Braumoeller 2003, 210). With the help of Maximum-Likelihood techniques, the impact of each causal mechanism (i.e. the vector of several independent variables) can then be tested. In other words, "[p]redictors influence the response variable singularly, and in combination with each other" (Porcu, Puggioni and Tedesco 2007, 196). This again shows the (non-additive) causal complexity that can be measured with the help of Boolean logit.

More formally, for each condition (or configuration/causal mechanism) A_k, a distinct latent dependent variable is assumed. Boolean logit then models the probabilities of each of these unobserved, i.e. latent, variables in a Boolean fashion as indicated in the model (either *and* or *or*). These conditions (A_k) then together or separately lead to the occurance of Y. Each of the conditions is determined by some vector of independent variables. The same independent variable may be included in different conditions A_k without any multicollinearity constraints (Porcu, Puggioni and Tedesco 2007).

One of the shortcomings of Boolean logit however is that the researcher needs to preliminarily posit a model (probability statements) and anticipate theoretically and subjectively how the conditions (A_k) are combined to lead to the occurance of Y (Braumoeller 2003; Hellström 2011). But with the help of previously run standard logit regressions and some descriptive statistics, hypotheses on the respective type(s) of causal complexity can be developed. Hence, Boolean logit is

> "neither an alternative nor a method better than the standard logistic one, but it does offer an advantage: it allows the researcher to consider models that consider causal complexity. […] [Hence,] … Boolean logit is a useful tool for implementing sensitivity analyses of other models and re-enforcing the evidence that emerges regarding the meaning of the predictors studied" (Porcu, Puggioni and Tedesco 2007, 205).

As such, Boolean logit results in the current stipend awarding context will be presented only in addition to the previously discussed standard logistic results.

Assumptions about multiple paths leading to selection success are derived from the results of Estimations I-XIII and will be discussed subsequently. Building on the previous results, it seems quite feasible that success is mainly dependent on previous academic performance measured in terms of high school and university grades. Applicants with poor grades are basically not awarded the stipend whereas better grades significantly increase selection probabilities. Consequently, a minimum academic performance seems necessary for selection success and a first condition or causal mechanism (A_1) leading to $Y_{Aw} = 1$ can be defined as 'academic performance'.

However, not everybody with grades better than X is subsequently awarded the stipend. This leads to the assumption that academic performance is only effective in combination with other signals. As has been shown in the (standard) logistic regressions, candidates who intensively pursued an ECA (high strength of ECAs) did have better chances to be awarded the scholarship than those who did not pursue any ECA at all. Thus, a second condition leading to selection success (A_2) can be defined as 'extracurricular activities'. But ECAs alone are not likely to explain somebody's success in selection either. Hence, pursuing an ECA intensively will only be a sufficient, but not a necessary condition and will only be effective in combination with sufficient academic performance (A_1).

Another way of (additionally) differentiating from the applicant pool is to be strongly recommended by somebody credible, for instance a professor. The corresponding causal mechanism or configuration (A_3) could be named 'recommendation'. This again will also not be promising per se, but is likely to only impact selection success in combination with sufficient academic performance.

Finally, it has been discovered that previous awards positively affect selection probabilities. Irrespective of the effects that drive the evaluator's preference for previous recipients of other (merit-based) sponsorships,[159] the mere fact of being previously awarded by another institution might be captured in a fourth configuration (A_4) named 'previous awards'. Similar to A_2 and A_3, this causal mechanism is anticipated to impact selection success only in combination with sufficient academic performance.

[159] As has been discussed in Chapter 7.3.1.4, the preferential treatment might either be explained by a mere Matthew Effect or by unobserved heterogeneity which is captured in the dummy variable 'Sponsored by the Studienstiftung des Deutschen Volkes'.

Consequently, a complex combination of the following form is likely to lead to selection success:

$$A_1 \cap (A_2 \cup A_3 \cup A_4)$$

Having excellent grades in this combination is a necessary condition for selection success, but not sufficient. Only in combination with either a high level of ECAs (A_2) or outstanding recommendations (A_3) or previous awards (A_4) $Y_{Aw} = 1$ is predicted to occur. The results of the Boolean logit modeling exactly this kind of complex combination are shown in table 7-22.

Boolean Logit Estimates **Dependent Variable: Scholarship Awarded (Yes=1 No=0)**		**Coef.**	**Std. Err.**	**z**	**P>|z|**
Path 1: Academic Performance (A_1)	High School Grade Average	**-1.144**	**0.302**	**-3.79**	**0.000**
	University Grade Average	**-1.752**	**0.355**	**-4.94**	**0.000**
	Constant	**-0.603**	**0.293**	**-2.06**	**0.039**
Path 2: Academic Performance and ECAs ($A_1 \cap A_2$)	One Type of ECAs	13.800	1204.86	0.01	0.991
	Constant	-14.374	1204.86	-0.01	0.990
Path 3: Academic Performance and Recommendation ($A_1 \cap A_3$)	Recommendation	**1.625**	**0.952**	**1.71**	**0.008**
	Recommender= Professor	0.959	0.855	1.12	0.262
	Good Relation to Recommending Person	0.665	0.982	0.68	0.498
	Constant	-0.918	1.155	-0.80	0.427
Path 4: Academic Performance and Previous Awards ($A_1 \cap A_4$)	'Studienstiftung des Deutschen Volkes'	128.158	.	.	.
	Applied for other Scholarships	13.697	714.177	0.02	0.985
	Constant	-13.330	714.178	-0.02	0.985
n				450	
Log-Likelihood				-214.06	
Wald Chi² (Prob> Chi²)				48.98 (0.000)	
(54 missing values generated)					
Correctly predicted 343 of 450 cases, or 76.22%					

Table 7-22 Boolean Logit Estimates Causal Complexity I

It appears from the Boolean logit estimates that the predicted causal complexity does not exist in the current case. Only academic performance as measured in terms of university and high school grades independently affects selection outcomes, but a joint influence of academic performance and ECAs (path 2) or academic performance and previous awards (path 4) could not be confirmed. Among the variables forming path 3, only the standardized recommendation (on a scale from 0 to 10) affects success in combination with academic performance significantly. All other influences disappear when combined with academic performance.

For the sake of completeness, other plausible causal connections of the above mentioned four conditions have also been tested and their results will be displayed below. It is possible, that all of the above mentioned conditions jointly make $Y_{Aw} = 1$ occur. This causal connection is represented as follows

$$A_1 \cap A_2 \cap A_3 \cap A_4$$

Results of this Boolean logit Model are shown in table 7-23. Again, this causal complexity cannot explain selection success sufficiently and only academic performance seems to drive selection outcomes.

Boolean Logit Estimates **Dependent Variable: Scholarship Awarded (Yes=1 No=0)**		**Coef.**	**Std. Err.**	**z**	**P>|z|**
Path 1: Academic Performance (A_1)	High School Grade Average	**-1.172**	**0.419**	**-3.51**	**0.000**
	University Grade Average	**-2.162**	**0.485**	**-4.46**	**0.000**
	Constant	-0.163	0.410	-0.40	0.691
Path 2: Academic Performance and ECAs ($A_1 \cap A_2$)	One Type of ECAs	-15.545	1807.63	-0.01	0.993
	Constant	17.331	1807.63	0.01	0.992
Path 3: Academic Performance, ECAs and Recommendation ($A_1 \cap A_2 \cap A_3$)	Recommendation	1.051	0.724	1.45	0.147
	Recommender= Professor	15.646	1647.16	0.01	0.992
	Good Relation to Recommending Person	16.357	1647.16	0.01	0.922
	Constant	-15.086	1647.16	-0.01	0.993
Path 4: Academic Performance, ECAs, Recommendation and Previous Awards ($A_1 \cap A_2 \cap A_3 \cap A_4$)	'Studienstiftung des Deutschen Volkes'	17.889	3869.94	0.00	0.996
	Applied for other Scholarships	0.621	0.539	1.15	0.250
	Constant	0.662	0.572	1.16	0.247
n				450	
Log-Likelihood				-212.71	
Wald Chi² (Prob> Chi²)				28.77 (0.000)	
(54 missing values generated)					
Correctly predicted 352 of 450 cases, or 78.22%					

Table 7-23 Boolean Logit Estimates Causal Complexity II

Finally, it is conceivable that applicants having achieved 'only' poor or average grades, could increase their chances to be selected by exhibiting a combination of strong ECA(s), excellent recommendation and previous awards. Hence, the Boolean logic behind this scenario would be

$$A_1 \cup (A_2 \cap A_3 \cap A_4)$$

The corresponding results for this model however once again show that only path 1, i.e. academic performance, influences selection outcomes significantly and consequently the results are not displayed here. In summary, only one of the anticipated multiple causal paths leading to selection success indeed affects selection outcomes, i.e. academic performance. Either academic performance in terms of grades is really the most important predictor of selection success, or more complex causal connections occur which cannot be modeled by Boolean logit.[160] Consequently, the Boolean logit results confirm the inconsistent interaction effects found in section 7.3.1.6 (the additive effect of different ability signals).

[160] The maximum number of causal paths which can be included in the Stata command *mlboolean* is four (Braumoeller 2004).

8 CONCLUSION

The present thesis is the first to systematically address scholarship selection processes in Germany. Previous research in this area is fragmentary at best and rather descriptive than analytical. Middendorff, Isserstedt and Kandulla (2009) for instance were the only researchers trying to find out how the 'average' stipend awardee looks like. With the help of a survey among actual scholarship holders, they adopted a questionnaire approach and e.g. found out that the proportion of upper class students is disproportionately high among awardees. This result corroborates the assertion that mainly already-privileged students benefit from merit-based scholarships. However, Middendorff, Isserstedt and Kandulla (2009) were only able to observe stipend awardee characteristics, but could not differentiate between supply- and demand-side effects. As a consequence, they could not deduce from their results whether e.g. upper class students do apply more frequently for such a scholarship (self-selection effects) or whether evaluators preferentially select these applicants (screening effects). In contrast to previous investigations, the present study aimed at clearly separating self-selection and screening effects. It is a first attempt to shed some light on factors of success of both written scholarship applications and face-to-face stipend awarding interviews. Using actual applications for a study-abroad scholarship offered by the DAAD, it was empirically tested which signals and indices influence evaluator decisions and selection outcomes.

Summary of Results

The empirical analysis of stipend selection processes has initially revealed that undergraduate students applying for a study-abroad scholarship do form a specific sub-group of all (undergraduate) students (who want to go abroad). To mention only a few differences, applicants do achieve significantly better grades – both in high school and university – and do engage more actively in extracurricular activities than the entire student body. Additionally, applicants attending higher education institutions in some specific German regions do apply more frequently than students from other German regions. Accordingly, applicants represent a positively self-selected group of all students in Germany. The decision to apply seems to be the result of an unobservable decision process made by every potential applicant. Presumably, potential applicants try to anticipate selection criteria that might be utilized by evaluators and do choose to apply only when they perceive their individual probability to be awarded the stipend to be sufficiently high.

Whenever an applicant decided to apply, the institutional decision process resulting in either success or failure could be examined. In the present investigation, it was found that the decisions are mainly based on rational considerations and consequently, applicants' ability signals predominantly affect selection outcomes. Ability signals that have been identified to be particularly influential are aspects connected with past performance (biodata) such as grades, recommendations and extracurricular activities. Intentions and expectations, e.g. in terms of desired guest institution type and quality, do only play a minor role.

It has also been found that signaling in the scholarship 'market' also works partly different than in the job market. Whereas grades are considered to have only a low validity to predict future job market success, they seem to be a valid and readily used signal in scholarship application processes.

Additionally, letters of reference – a selection tool which is undervalued in hiring decisions – play an important role in scholarship applications. Especially, the status of the recommending person (professor) helps the applicant to pass the first hurdle in a selection process: Evaluators might be impressed by the student's ability to have already established a good reputation at the very beginning of their studies. An undergraduate student already staying in close contact with a professor might be instantly perceived as being an excellent student.

The consideration of family background and/or part-time employment in pre-selection indicates a certain social aspect (or: positive discrimination) to scholarship awarding: not only the best in terms of academic achievement, but also those who seem to be more in (financial) need are preferentially selected as applicants indicating their parents to have an academic background are less frequently invited to an interview. Consequently, the assumption that stipends are predominantly awarded to students with upper class background was not supported and the results of Middendorff, Isserstedt and Kandulla (2009) seem to represent supply-side effects (self-selection) only. Hence, upper class students might be more inclined to apply for a merit-based scholarship than (otherwise similar) middle or working class students, but upon condition that both decided to apply, evaluators c.p. do not choose the upper class applicant over the working class applicant, but the other way round.

A positive effect of being previously elected by another scholarship granting institution (especially the 'Studienstiftung des Deutschen Volkes') could be found in both pre- and final selection: Those already being supported by another institution seem to benefit both in paper applications and during selection interviews. Whether their specific advantage arises from a mere Matthew Effect, their general conversance with selection interviews or whether they possess some other qualifications (e.g. eloquence or personality) that could not be considered in this study (unobserved heterogeneity), could not be determined in this study and still needs to be investigated in future research.

Evidence for discriminatory behavior based on applicant individual difference factors was only found during pre-selection. Everything else being equal, applicants who were born outside of Germany face an inherent disadvantage. Although only 41 of all 504 applicants were not born in Germany, the reasons for their 'discrimination' still need to be examined. As having a German high school diploma and studying at a German higher education institution are prerequisites for applying in the investigated program, there should be no doubt about the appropriateness of these applicants' academic qualification. As diverse qualified information is made available in the application, information asymmetries should usually be small and statistical discrimination is not likely to occur. Therefore, only taste-based discrimination might be the rationale behind placing these applicants at a competitive disadvantage. Alternatively, unobserved heterogeneity could serve as an explanation for the less favorable pre-selection treatment of applicants who were not born in Germany: It is possible that this group of applicants accidentally differs from all other applicants in some characteristics that we were not able to observe or measure in the current project. The assumption that other unobserved variables than pure discrimination drive the 'Not-born-in-Germany' effect is reinforced by the fact that neither having a foreign-sounding name nor having a foreign citizenship is associated with significantly worse awarding probabilities, but solely an applicant's country of birth has a significant negative impact on awarding probabilities.

In addition to applicant individual difference factors, also some evaluator characteristics have been shown to lead to different selection outcomes. In line with previous findings from personnel selection, female pre-selection evaluators proved to be more lenient when it comes to interview invitations (regardless of applicant gender). In panel interview situations however, the fraction of female evaluators in the panel does not affect selection outcomes (neither for female nor for male applicants).

During final selection interviews, one of the few significant (and testable) factors of influence was interview time. Applicants being interviewed between 11:00 and 11:59 am do have significantly lower chances of being awarded the stipend (everything else being equal). This competitive disadvantage might be explained by a combination of evaluator behavior and the structural framework of decision making. Listening to interviews from 9 a.m. until 7 p.m. with only a short lunch break from 1 to 2 pm might be an exhausting task leading to especially rigorous decisions during certain spots in the middle of the day. If indeed evaluator fatigue is the explanation for differential success probabilities, an increased interview structure as presented in the literature review might serve as a remedy for this extraneous (and unsolicited) effect. Reliable statements about the rationale behind this factor of influence however can only be made after altering (i.e. restructuring) the final selection process and replicating the present investigation.

To sum up, success factors of applications differ widely in pre- and final selection. Whereas signals provided in written scholarship applications (e.g. grades, language tests and extracurricular activities) have been revealed to be extremely effective in scholarship pre-selection and are able to explain a high portion of variation in pre-selection outcomes, most of these influences (except for university grades) become obsolete as soon as the applicant gets the opportunity to present herself in a face-to-face interview. This closely corresponds to the P-J and P-O fit constructs in personnel selection: Recruiters sift applications on the basis of applicants' KSAs and do only invite candidates with sufficient P-J fit to an interview. Subsequently, final selection is based on the level of P-O fit, i.e. mainly on applicant characteristics that could not be revealed from paper applications, but are only observable in face-to-face interviews.

In the present case of scholarship applications, interview success also seems to be driven by certain unobserved variables and an omitted variable bias is very likely to occur. Part of the unobserved heterogeneity may be explained by applicant impression management, eloquence, personality or presentation skills as these have been found to influence interview success in a series of other empirical studies. All of these factors however could not be tested with the available applicant information and would have entailed further complex (and expensive) data collection. Overall, social factors such as applicant-rater similarity (at least in terms of gender, regional and institutional similarity) do not seem to bias the selection outcome in scholarship applications, neither in pre- nor in final selection decisions. Unlike recruiters, stipend awarding evaluators do not treat applicants more favorably just on the basis of (demographic or attitudinal) similarity. This result illustrates vividly that recruiters and stipend awarding evaluators follow

distinct utility functions. The absence of any similarity-attraction effect in stipend awarding decisions might be explained by a certain emotional distance between rater and applicant. As future contact intensity between evaluator and stipend awardee is very low (if non-existent), personal similarity does only play a minor part in scholarship selection decisions.

All in all, decision processes in stipend awarding decisions seem to be much more rational and especially based on applicant ability signals than personnel selection decisions.

Relation to Theory

From a sociological perspective, it was hypothesized that (merit-based) scholarship granting organizations do aim at selecting future elite members. This objective has explicitly been stated by the German Federal Ministry of Education and Research (BMBF 2009). Consequently, several elite theories were consulted for theoretically explaining stipend awarding decisions. In line with functional elite theory, an applicant's field of study did not affect stipend awarding success in the investigated selection process. As a result, access to stipends is equally gained to students from all fields. Assuming that the investigated stipend facilitates access to elite positions, the notion of several parallel functional elites is indeed supported by the empirical results. Additionally, the existence of performance elites is supported by the previously discussed results as performance is the main driver of stipend awarding decisions.

From an economic perspective on the other hand, the selection process is interpreted as an investment decision under uncertainty. Based on agency theory, the principal (here: the evaluator) is not able to observe the true quality of the agent (here: the applicant) as information asymmetries occur. As a result, agents need to signal their ability and principals need to screen their applications. Economically rational decisions are then entirely based on ability signals and inefficiencies such as discrimination based on applicant (or rater) individual difference factors do not occur. The present study has revealed that stipend awarding decisions are based on ability signals to a great extent. Only very few indices (e.g. country of birth) do have an impact on awarding decisions.

Limitations and Future Research

The present thesis was the first to systematically address the scholarship selection process by empirically testing which signals and indices influence evaluator decisions and consequently, selection outcomes.

Due to the rather small number of observations (429 usable application sets in pre-selection and 226 in final selection) and the specificity of the investigated program however, the results need to be reinforced by other investigations of a similar kind. In order to obtain more general results, future studies also need to examine selection processes in countries other than Germany. Moreover, future research should focus more on certain variables that revealed to be of importance in the current study (e.g. grades, place of birth, extracurricular activities). Field experiments (or even correspondence studies) could actively manipulate certain signals and indices and measure their respective effect more precisely.

Furthermore, the present study's results may be influenced by the specific requirements of a stay abroad (language skills for instance). Subsequent research should additionally address other scholarship purposes in order to find more general success factors in scholarship applications. Likewise, different selection methods and their respective influence should be investigated in future research projects. Particularly interesting could be whether or not the application success factors differ when other evaluators than professors decide upon selection or rejection.

Nonetheless, the present study is able to provide (education) economists and policy advisors with a better understanding of scholarship selection processes and the nature and composition of student elites arising from such academic distinctions. The empirical method developed for and applied in the present thesis can be understood as some kind of controlling mechanism for stipend awarding decisions in general. As a consequence, it can be easily transferred to stipend awarding processes other than the ones observed at the DAAD and can serve as a powerful to control for the degree of stipend goal attainment.

With reference to the initially presented objective of counteracting the imminent long-term shortage of skilled labor in Germany by offering more merit-based scholarships, other research questions emerge which could not be answered with the available research design.

One of these important questions is to find out whether the ambitious goal stated by the BMBF (and in the DAAD's mission statement) is accomplished: Do merit-based scholarships really attract the most promising students? In other words, do stipend awardees really turn out to be future elite members? This however can only be studied in a longitudinal follow-up research design illustrating later life performance of stipend awardees. Only by tracking the long-term career-path of stipend awardees, one might be able to find out whether or not evaluators chose the 'right' applicants. Although the present thesis is not able to answer all the questions connected with stipend awarding decisions, it has laid the foundation of an innovative, promising and socially relevant strand of research that requires constant attention in future.

REFERENCES

AAMODT, M.G.; BRYAN, D.A.; WHITCOMB, A.J. (1993). Predicting Performance with Letters of Recommendation, *Public Personnel Management,* 22, 81-90.

ABREVAYA, J.; HAMERMESH, D.S. (2010). Charity and Favoritism in the Field – Are Female Economists Nicer (To Each Other)? *IZA Working Paper* No. 4921.

ADAMS, G.R. (1977). Physical Attractiveness Research – Toward a Developmental Social Psychology of Beauty, *Human Development*, 20, 217-239.

ADKINS, C.L.; RUSSELL, C.J.; WERBEL, J.D. (1994). Judgments of Fit in the Selection Process – The Role of Work Value Congruence, *Personnel Psychology*, 47, 605-623.

AGTHE, M.; SPÖRRLE, M.; MANER, J.K. (2010). Don't Hate Me Because I'm Beautiful – Anti-Attractiveness Bias in Organizational Evaluation and Decision Making, *Journal of Experimental Social Psychology,* 46, 1151-1154.

AHMED, A.M.; ANDERSSON, L.; HAMMARSTEDT, M. (2012). Does Age Matter for Employability? A Field Experiment on Ageism in the Swedish Labour Market, *Applied Economics Letters*, 19, 403-406.

AHMED, A.M.; ANDERSSON, L.; HAMMARSTEDT, M. (2009). Ethnic Discrimination in the Market Place of Small Business Transfers, *Economics Bulletin*, 29, 3050-3058.

AI, C.; NORTON, E.C. (2003). Interaction Terms in Logit and Probit Models, *Economics Letters*, 80, 123-129.

AKAIKE, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle, in Petrov, B.N.; Csaki, B.F. (Editors): Second International Symposium on Information Theory, Academiai Kiado, Budapest, 267-281.

AKERLOF, G. (1970). The Market for "Lemons" – Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics*, 84, 488-500.

ALBERT, R.; ESCOT, L.; FERNÁNDEZ-CORNEJO, J.A. (2011). A Field Experiment to Study Sex and Age Discrimination in the Madrid Labour Market, *The International Journal of Human Resource Management*, 22, 351-375.

ANDREONI, J.; VESTERLUND, L. (2001). Which Is the Fair Sex? Gender Differences in Altruism, *Quarterly Journal of Economics*, 116, 293-312.

ANWAR, S.; BAYER, P.; HJALMARSSON, R. (2012). A Fair and Impartial Jury? The Role of Age in Jury Selection and Trial Outcome, *NBER Working Paper* No. 17887.

ARGYLE, M.; MCHENRY, R. (1971). Do Spectacles Really Affect Judgements of Intelligence? *British Journal of Social and Clinical Psychology*, 10, 27-29.

ARROW, K.J. (1973). The Theory of Discrimination, in Ashenfelter, O.; Rees, A. (Editors): Discrimination in Labor Markets, Princeton University Press, Princeton, NJ, 3-33.

ASCH, S.E. (1956). Studies of Independence and Conformity – I. A Minority of One against a Unanimous Majority, *Psychological Monographs: General and Applied*, 70, 1-70.

ATHEY, S.; KATZ, L.F.; KRUEGER, A.B.; LEVITT, S.; POTERBA, J. (2007). What Does Performance in Graduate School Predict? Graduate Economics Education and Student Outcomes, *American Economic Review: Papers and Proceedings,* 97, 512-518.

BACKHAUS, K.; ERICHSON, B.; PLINKE, W.; WEIBER, R. (2011). Multivariate Analysemethoden – Eine anwendungsorientierte Einführung, 13th Ed., Springer-Verlag, Heidelberg.

BAGÜES, M.F.; ESTEVE-VOLART, B. (2010). Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment, *Review of Economic Studies,* 77, 1301-1328.

BAHNS, A.J.; PICKETT, K.M.; CRANDALL, C.S. (2012). Social Ecology of Similarity – Big Schools, Small Schools and Social Relationships, *Group Processes & Intergroup Relations*, 15, 119-131.

BARRICK, M.R.; MOUNT, M.K. (1991). The Big Five Personality Dimensions and Job Performance – A Meta-Analysis, *Personnel Psychology*, 44, 1-26.

BARRICK, M.R.; MOUNT, M.K.; JUDGE, T.A. (2001). Personality and Performance at the Beginning of the New Millennium – What Do We Know and Where Do We Go Next? *International Journal of Selection and Assessment*, 9, 9-30.

BARRICK, M.R.; PATTON, G.K.; HAUGLAND, S.N. (2000). Accuracy of Interviewer Judgments of Job Applicant Personality Traits, *Personnel Psychology*, 53, 925-951.

BARRICK, M.R.; SHAFFER, J.A.; DEGRASSI, S.W. (2009). What You See May Not Be What You Get – Relationships Among Self-Presentation Tactics and Ratings of Interview and Job Performance, *Journal of Applied Psychology*, 94, 1394-1411.

BAXTER, J.C.; BROCK, B.; HILL, P.C.; ROZELLE, R.M. (1981). Letters of Recommendation – A Question of Value, *Journal of Applied Psychology,* 66, 296-301.

BECKER, G.S. (1993). Human Capital – A Theoretical and Empirical Analysis, with Special Reference to Education, 3rd Ed., The University of Chicago Press, Chicago, IL.

BECKER, G.S. (1971). The Economics of Discrimination, 2nd Ed., The University of Chicago Press, Chicago, IL.

BECTON, J.B.; MATTHEWS, M.C.; HARTLEY, D.L.; WHITAKER, D.H. (2009). Using Biodata to Predict Turnover, Organizational Commitment, and Job Performance in Healthcare, *International Journal of Selection and Assessment,* 17, 189-202.

BEEHR, T.A.; GILMORE, D.C. (1982). Applicant Attractiveness as a Perceived Job-Relevant Variable in Selection of Management Trainees, *The Academy of Management Journal*, 25, 607-617.

BEHRENZ, L. (2001). Who Gets the Job and Why? An Explorative Study of Employers' Recruitment Behavior, *Journal of Applied Economics,* 4, 255-278.

BENDICK JR., M. (2007). Situation Testing for Employment Discrimination in the United States of America, *Horizons stratégiques,* 3/2007, 17-39.

BENDICK JR., M. (1983). The Role of Public Programs and Private Markets in Reemploying Displaced Workers, *Policy Studies Review,* 2, 715-733.

BENDICK JR., M.; BROWN, L.E.; WALL, K. (1999). No Foot in the Door – An Experimental Study of Employment Discrimination against Older Workers, *Journal of Aging & Social Policy*, 10, 5-23.

BENDICK JR., M.; JACKSON, C.W.; REINOSO, V.A. (1994). Measuring Employment Discrimination through Controlled Experiments, *Review of Black Political Economy,* 23, 25-48.

BENDICK JR., M.; JACKSON, C.W.; REINOSO, V.A.; HODGES; L.E. (1991). Discrimination against Latino Job Applicants – A Controlled Experiment, *Human Resource Management,* 30, 469-484.

BENDICK JR., M.; JACKSON, C.W.; ROMERO, J.H. (1996). Employment Discrimination against Older Workers – An Experimental Study of Hiring Practices, *Journal of Aging & Social Policy*, 8, 25-46.

BERRY, C.M.; GRUYS, M.L.; SACKETT, P.R. (2006). Educational Attainment as a Proxy for Cognitive Ability in Selection – Effects on Levels of Cognitive Ability and Adverse Impact, *Journal of Applied Psychology,* 91, 696-705.

BERSCHEID, E.; WALSTER, E. H. (1969). Interpersonal Attraction, Addison-Wesley, Reading, MA.

BERTELSMANN STIFTUNG (2012). Chancenspiegel – Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme – Zusammenfassung zentraler Befunde – Eine Studie des Instituts für Schulentwicklungsforschung (IFS) der Technischen Universität Dortmund in Kooperation mit der Bertelsmann Stiftung, Verlag Bertelsmann Stiftung, Gütersloh, accessed October 24, 2012 at: www.chancen-spiegel.de.

BERTRAND, M.; MULLAINATHAN, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination, *American Economic Review,* 94, 991-1013.

BIDDLE, J.E.; HAMERMESH, D.S. (1998). Beauty, Productivity, and Discrimination – Lawyers' Looks and Lucre, *Journal of Labor Economics*, 16, 172-201.

BLINDER, A. (1973). Wage Discrimination – Reduced Form and Structural Estimates, *The Journal of Human Resources, 8*, 436-455.

BMBF/BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2012a). Informationen zum Deutschlandstipendium, Bonn and Berlin, accessed October 24, 2012 at: http://deutschlandstipendium.de/_media/flyer_deutschlandstipendium.pdf.

BMBF/BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2012b). Stipendienlotse – Der Online-Wegweiser für Stipendien, accessed October 24, 2012 at: http://www.bmbf.de/pub/flyer_stipendienlotse.pdf.

BMBF/BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2011). Mehr als ein Stipendium – Staatliche Begabtenförderung im Hochschulbereich, Bonn and Berlin, accessed October 24, 2012 at: http://www.bmbf.de/pub/flyer_staatliche_begabtenfoerderung_hochschulbereich.pdf

BMBF/BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2010). Das neue BAföG – Informationen zur Ausbildungsförderung, Bonn and Berlin, accessed October 24, 2012 at: http://www.bmbf.de/pub/bafoeg_flyer.pdf.

BMBF/BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2009). Mehr als ein Stipendium – Staatliche Begabtenförderung im Hochschulbereich, 1. Aktualisierte Auflage, Bonn and Berlin, accessed October 24, 2012 at: http://www.stipendiumplus.de/_media/BMBF_Begabten_Broschuere0409.pdf

BOLINO, M.C.; KACMAR, K.M.; TURNLEY, W.H.; GILSTRAP, J.B. (2008). A Multi-Level Review of Impression Management Motives and Behaviors, *Journal of Management*, 34, 1080-1109.

BOOTH, A.; LEIGH, A. (2010). Do Employers Discriminate by Gender? A Field Experiment in Female-Dominated Occupations, *IZA Discussion Paper* No. 4690.

BORJAS, G.J.; GOLDBERG, M.S. (1978). Biased Screening and Discrimination in the Labor Market, *American Economic Review,* 68, 918-922.

BOSHIER, R. (1975). A Video-Tape Study on the Relationship between Wearing Spectacles and Judgments of Intelligence, *Perceptual and Motor Skills*, 40, 69-70.

BOUDREAU, J.W.; BOSWELL, W.R.; JUDGE, T.A.; BRETZ JR, R.D. (2001). Personality and Cognitive Ability as Predictors of Job Search among Employed Managers, *Personnel Psychology*, 54, 25-50.

BOURDIEU, P. (1996). Über die Beziehungen zwischen Geschichte und Soziologie in Frankreich und Deutschland – Pierre Bourdieu im Gespräch mit Lutz Raphael, *Geschichte und Gesellschaft*, 22, 62-89.

BOURDIEU, P. (1991). Das Feld der Macht und die technokratische Herrschaft, in Dölling, I. (Editor): Die Intellektuellen und die Macht, VSA Verlag, Hamburg, 67-100.

BOURDIEU, P. (1989). La Noblesse d'État – Grandes Écoles et Esprit de Corps, Les Éditions de Minuit, Paris.

BOURDIEU, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital, in Kreckel, R. (Editor): Soziale Ungleichheiten, Soziale Welt Sonderband 2, Verlag Otto Schwarz & Co, Göttingen, 183-198.

BOVENKERK, F.; GRAS, M.J.I.; RAMSOEDH, D.; WITH THE ASSISTANCE OF DANKOOR, M.; HAVELAAR, A. (1995). Discrimination against Migrant Workers and Ethnic Minorities in Access to Employment in the Netherlands, *International Migration Papers* No. 4, International Labour Office, Geneva, Switzerland.

BRAUMOELLER, B.F. (2004). Boolean Logit and Probit in Stata, *The Stata Journal,* 4, 436-441.

BRAUMOELLER, B.F. (2003). Causal Complexity and the Study of Politics, *Political Analysis,* 11, 209-233.

BRENKE, K. (2010). Fachkräftemangel kurzfristig noch nicht in Sicht, *Wochenbericht des Deutschen Instituts für Wirtschaftsforschung (DIW)*, 77 (46), 2-15.

BRETZ JR., R.D.; RYNES, S.L.; GERHART, B. (1993). Recruiter Perceptions of Applicant Fit – Implications for Individual Career Preparation and Job Search Behavior, *Journal of Vocational Behavior*, 43, 310-327.

BROCKHAUS (2006). Brockhaus Enzyklopädie in 30 Bänden – Band 7 DIEU-EMAR, 21st Ed., F.A. Brockhaus, Leipzig/Mannheim.

BRODER, I.E. (1993). Review of NSF Economics Proposals – Gender and Institutional Patterns, *American Economic Review*, 83, 964-970.

BROWN, B.K.; CAMPION, M.A. (1994). Biodata Phenomenology – Recruiters' Perceptions and Use of Biographical Information in Résumé Screening, *Journal of Applied Psychology*, 79, 897-908.

BROWN, M.J.; HENRIQUEZ, E.; GROSCUP, J. (2008). The Effects of Eyeglasses and Race on Juror Decisions Involving a Violent Crime, *American Journal of Forensic Psychology*, 26, 25-43.

BUCHE, A.; CARSTENSEN, J. (2009). Qualitative Comparative Analysis – Ein Überblick, in Kriwy, P.; Gross, C. (Editors): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen, VS Verlag für Sozialwissenschaften, Wiesbaden, 65-92.

BUCKLEY, M.R.; JACKSON, K.A.; BOLINO, M.C.; VERES, J.G.; FIELD, H.S. (2007). The Influence of Relational Demography on Panel Interview Ratings – A Field Experiment, *Personnel Psychology,* 60, 627-646.

BUNDESAGENTUR FÜR ARBEIT (2011). Hintergrundinformation – Aktuelle Fachkräfteengpässe, Arbeitsmarktberichterstattung, Nürnberg, accessed October 24, 2012 at: http://statistik.arbeitsagentur.de/Statischer-Content/Arbeitsmarktberichte/Berichte-Broschueren/Arbeitsmarkt/Generische-Publikationen/Hintergrundinfo-Fachkraeftebedarf-2011-04.pdf

BURSELL, M. (2007). What's in a Name? A Field Experiment Test for the Existence of Ethnic Discrimination in the Hiring Process, *SULCIS Working Paper* 2007:7.

BÜRKLIN, W.; REBENSTORF, H. ET AL. (1997). Eliten in Deutschland – Rekrutierung und Integration, Leske + Budrich, Opladen.

BÜSCH, V.; DAHL, S.-A.; DITTRICH, D.A.V. (2009). An Empirical Study of Age Discrimination in Norway and Germany, *Applied Economics*, 41, 633-651.

BUSINESS DICTIONARY (2012). Stipend, accessed October 24, 2012 at: http://www.businessdictionary.com/definition/stipend.html.

BUSS, D.M. (1985). Human Mate Selection, *American Scientist*, 73, 47-51.

BYRNE, D. (1971). The Attraction Paradigm, Academic Press, New York, NY.

BYRNE, D.; NEUMANN, J. H. (1992). The Implications of Attraction Research for Organizational Issues, *Advances in Psychology, 82*, 29-70.

CABLE, D.M.; DERUE, D.S. (2002). The Convergent and Discriminant Validity of Subjective Fit Perceptions, *Journal of Applied Psychology*, 87, 875-884.

CABLE, D.M.; GILOVICH, T. (1998). Looked Over or Overlooked? Prescreening Decisions and Postinterview Evaluations, *Journal of Applied Psychology*, 83, 501-508.

CABLE, D.M.; JUDGE, T.A. (1997). Interviewers' Perceptions of Person-Organization Fit and Organizational Selection Decisions, *Journal of Applied Psychology*, 82, 546-561.

CAIN, G.C. (1986). The Economic Analysis of Labor Market Discrimination – A Survey, in Ashenfelter, O.; Layard, R. (Editors): Handbook of Labor Economics, Volume I, Elsevier North Holland, Amsterdam, 693-785.

CALDWELL, D.F.; BURGER, J.M. (1998). Personality Characteristics of Job Applicants and Success in Screening Interviews, *Personnel Psychology*, 51, 119-136.

CALLANAN, G.; BENZIG, C. (2004). Assessing the Role of Internships in the Career-Oriented Employment of Graduating College Students, *Education + Training, 46*, 82-89.

CAMPION, M.A. (1978). Identification of Variables most Influential in Determining Interviewers' Evaluations of Applicants in a College Placement Center, *Psychological Reports*, *42*, 947-952.

CAMPION, M.A.; PALMER, D.K.; CAMPION, J.E. (1997). A Review of Structure in the Selection Interview, *Personnel Psychology*, 50, 655-702.

CANN, A.; SIEGFRIED, W.D.; PEARCE, L. (1981). Forced Attention to Specific Applicant Qualifications – Impact on Physical Attractiveness and Sex of Applicant Biases, *Personnel Psychology*, 34, 65-75.

CARLSSON, M. (2010). Experimental Evidence of Discrimination in the Hiring of First- and Second-Generation Immigrants, *Labour, 24*, 263-278.

CARLSSON, M.; ROOTH, D.-O. (2008a). Is It Your Foreign Name or Foreign Qualifications? An Experimental Study of Ethnic Discrimination in Hiring, *IZA Discussion Paper* No. 3810.

CARLSSON, M.; ROOTH, D.-O. (2008b). An Experimental Study of Sex Segregation in the Swedish Labour Market – Is Discrimination the Explanation? *IZA Discussion Paper* No. 3811.

CARLSSON, M.; ROOTH, D.-O. (2007). Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data, *Labour Economics, 14*, 716-729.

CARR, S. E. (2009). Emotional Intelligence in Medical Students – Does It Correlate with Selection Measures? *Medical Education, 43*, 1069-1077.

CAWLEY, J. (2004). The Impact of Obesity on Wages, *Journal of Human Resources*, 39, 451-474.

CHAPMAN, D.S.; ROWE, P.M. (2001). The Impact of Videoconference Technology, Interview Structure, and Interviewer Gender on Interviewer Evaluations in the Employment Interview – A Field Experiment, *Journal of Occupational and Organizational Psychology*, 74, 279-298.

CHATMAN, J.A. (1991). Matching People and Organizations – Selection and Socialization in Public Accounting Firms, *Administrative Science Quarterly*, 36, 459-484.

CHEN, C.-C.; YANG, I. W.-F.; LIN, W.-C. (2010). Applicant Impression Management in Job Interview – The Moderating Role of Interviewer Affectivity, *Journal of Occupational and Organizational Psychology*, 83, 739-757.

CHIA, Y.M. (2005). Job Offers of Multi-National Accounting Firm – The Effects of Emotional Intelligence, Extra-Curricular Activities, and Academic Performance, *Accounting Education: An international journal, 14*, 75-93.

CLEVELAND, J.N.; LANDY, F.J. (1981). The Influence of Rater and Ratee Age on Two Performance Judgments, *Personnel Psychology*, 34, 19-29.

CLIFFORD, M.M.; WALSTER, E. (1973). The Effect of Physical Attractiveness on Teacher Expectation, *Sociology of Education,* 46, 248-258.

COHEN, P.A. (1984). College Grades and Adult Achievement – A Research Synthesis, *Research in Higher Education,* 20, 281-293.

COLARELLI, S.M.; HECHANOVA-ALAMPAY, R.; CANALI, K.G. (2002). Letters of Recommendation – An Evolutionary Psychological Perspective, *Human Relations,* 55, 315-344.

COLE, M.S.; FEILD, H.S.; GILES, W.F. (2003a). Using Recruiter Assessments of Applicants' Résumé Content to Predict Applicant Mental Ability and Big Five Personality Dimensions, *International Journal of Selection and Assessment*, 11, 78-88.

COLE, M.S.; FEILD, H.S.; GILES, W.F. (2003b). What Can We Uncover about Applicants Based on Their Résumés? A Field Study, *Applied HRM Research*, 8, 51-62.

COLE, M.S.; FEILD, H.S.; GILES; W.F.; HARRIS, S.G. (2009). Recruiters' Inferences of Applicant Personality Based on Résumé Screening – Do Paper People Have a Personality? *Journal of Business Psychology,* 24, 5-18.

COLE, M.S.; FEILD, H.S.; GILES; W.F.; HARRIS, S.G. (2004). Job Type and Recruiters' Inferences of Applicant Personality Drawn from Résumé Biodata – Their Relationships with Hiring Recommendations, *International Journal of Selection and Assessment*, 12, 363-367.

COLE, M.S.; FEILD, H.S.; STAFFORD, J.O. (2005). Validity of Résumé Reviewers' Inferences Concerning Applicant Personality Based on Résumé Evaluation, *International Journal of Selection and Assessment*, 13, 321-324.

COLE, M.S.; RUBIN, R.S.; FEILD, H.S.; GILES, W.F. (2007). Recruiters' Perceptions and Use of Applicant Résumé Information – Screening the Recent Graduate, *Applied Psychology: An International Review*, 56, 319-343.

CONLEY, D.; GLAUBER, R. (2005). Gender, Body Mass and Economic Status, *NBER Working Paper* No. 11343.

COOK, M. (2009). Personnel Selection – Adding Value Trough People, 5th Ed., Wiley-Blackwell, Chichester.

COOK, K.W.; VANCE, C.A.; SPECTOR, P.E. (2000). The Relation of Candidate Personality With Selection-Interview Outcomes, *Journal of Applied Social Psychology*, 30, 867-885.

CORNWELL, C.; MUSTARD, D.B.; SRIDHAR, D.J. (2006). The Enrollment Effects of Merit-Based Financial Aid – Evidence from Georgia's HOPE Program, *Journal of Labor Economics*, 24, 761-786.

COSTA, P.T.; MCCRAE, R.R. (1997). Longitudinal Stability of Adult Personality, in Hogan, R.; Johnson, J. A.; Briggs, S.R. (Editors): Handbook of Personality Psychology, Academic Press, San Diego, CA, 269-290.

COSTA, P.T.; MCCRAE, R.R. (1992). The NEO PI-R Professional Manual, Psychological Assessment Resources, Odessa, FL.

COX, C.L.; GLICK, W.H. (1986). Résumé Evaluations and Cosmetics Use – When More Is Not Better, *Sex Roles*, 14, 51-58.

COX, D.R.; SNELL, E.J. (1989). Analysis of Binary Data, 2nd Ed., Chapman and Hall, London.

CROSON, R.; GNEEZY, U. (2009). Gender Differences in Preferences, *Journal of Economic Literature*, 47, 448-474.

CROSS, H.; KENNEY, G.M..; MELL, J.; ZIMMERMANN, W. (1990). Employer Hiring Practices – Differential Treatment of Hispanic and Anglo Job Seekers, The Urban Institute Press, Washington, DC.

DAHRENDORF, R. (1965). Gesellschaft und Demokratie in Deutschland, R. Piper & Co, München.

DAHRENDORF, R. (1962). Eine neue deutsche Oberschicht? Notizen über die Eliten in der Bundesrepublik, *Die neue Gesellschaft*, 9, 18-31.

DALESSIO, A.; IMADA, A.S. (1984). Relationships between Interview Selection Decisions and Perceptions of Applicant Similarity to an Ideal Employee and Self – A Field Study, *Human Relations*, 37, 67-80.

DANIEL, W.W. (1968). Racial Discrimination in England – Based on the P.E.P. Report, Penguin Books, Middlesex.

DANIEL, K.; HEYWOOD, J.S. (2007). The Determinants of Hiring Older Workers – UK Evidence, *Labour Economics*, 14, 35-51.

DANY, F.; TORCHY, V. (1994). Recruitment and Selection in Europe – Policies, Practices and Methods, in Brewster, C.; Hegewisch, A. (Editors): Policy and Practice in European Human Resource Management – The Price Waterhouse Cranfield Survey, Routledge, London, 68-88.

DANZIGER, S.; LEVAV, J.; AVNAIM-PESSO, L. (2011). Extraneous Factors in Judicial Decisions, *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 108, 6889-6892.

DAVISON, H.K.; BURKE, M.J. (2000). Sex Discrimination in Simulated Employment Contexts – A Meta-Analytic Investigation, *Journal of Vocational Behavior*, 56, 225-248.

DEAUX, K.; FARRIS, E. (1975). Complexity, Extremity, and Affect in Male and Female Judgments, *Journal of Personality*, 43, 379-389.

DE FRUYT, F.; MERVIELDE, I. (1999). RIASEC Types and Big Five Traits as Predictors of Employment Status and Nature of Employment, *Personnel Psychology*, 52, 701-727.

DELERY, J.E.; KACMAR, K.M. (1998). The Influence of Applicant and Interviewer Characteristics on the Use of Impression Management, *Journal of Applied Social Psychology*, 28, 1649-1669.

DE PRADA, M.A.; ACTIS, W.; PEREDA, C.; MOLINA, R.P. (1996). Labour Market Discrimination against Migrant Workers in Spain, *International Migration Papers* No. 9, International Labour Office, Geneva.

DESTATIS/STATISTISCHES BUNDESAMT (2012a). Bevölkerung und Erwerbstätigkeit – Bevölkerungsfortschreibung 2010, Fachserie 1, Reihe 1.3, Statistisches Bundesamt, Wiesbaden.

DESTATIS/STATISTISCHES BUNDESAMT (2012b). Hochschulen auf einen Blick – Ausgabe 2012, Statistisches Bundesamt, Wiesbaden, accessed October 24, 2012 at: https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/BroschuereHochschulenBlick0110010127004.pdf?__blob=publicationFile.

DESTATIS/STATISTISCHES BUNDESAMT (2012c). Bildung und Kultur – Personal an Hochschulen 2011, Fachserie 11, Reihe 4.4, Statistisches Bundesamt, Wiesbaden.

DESTATIS/STATISTISCHES BUNDESAMT (2011a). Bildungsstand der Bevölkerung 2011, Statistisches Bundesamt, Wiesbaden.

DESTATIS/STATISTISCHES BUNDESAMT (2011b). Bildung und Kultur – Prüfungen an Hochschulen 2010, Fachserie 11, Reihe 4.2, Statistisches Bundesamt, Wiesbaden.

DESTATIS/STATISTISCHES BUNDESAMT (2011c). Bildung und Kultur – Ausbildungsförderung nach dem Bundesausbildungsförderungsgesetz (BAföG) 2010, Fachserie 11, Reihe 7, Statistisches Bundesamt, Wiesbaden.

DESTATIS/STATISTISCHES BUNDESAMT (2009). Bildung und Kultur – Studierende an Hochschulen Wintersemester 2008/2009, Fachserie 11, Reihe 4.1, Statistisches Bundesamt, Wiesbaden.

DESTATIS/STATISTISCHES BUNDESAMT (2008). Bildung und Kultur – Studierende an Hochschulen Wintersemester 2007/2008, Fachserie 11, Reihe 4.1, Statistisches Bundesamt, Wiesbaden.

DIGMAN, J.M. (1990). Personality Structure – Emergence of the Five-Factor Model, *Annual Review of Psychology*, 41, 417-440.

DION, K; BERSCHEID, E.; WALSTER, E. (1972). What Is Beautiful Is Good, *Journal of Personality and Social Psychology*, 24, 285-290.

DIPBOYE, R.L. (1982). Self-Fulfilling Prophecies in the Selection/Recruitment Interview, *Academy of Management Review*, 7, 579-587.

DIPBOYE, R.L.; ARVEY, R.D..; TERPSTRA, D.E. (1977). Sex and Physical Attractiveness of Raters and Applicants as Determinants of Résumé Evaluations, *Journal of Applied Psychology,* 62, 288-294.

DIPBOYE, R.L.; FROMKIN, H.L.; WIBACK, K. (1975). Relative Importance of Applicant Sex, Attractiveness, and Scholastic Standing in Evaluation of Job Applicant Résumés, *Journal of Applied Psychology,* 60, 39-43.

DIXON, M; WANG, S.; CALVIN, J.; DINEEN, B.; TOMLINSON, E. (2002). The Panel Interview – A Review of Empirical Research and Guidelines for Practice, *Public Personnel Management*, 31, 397-428.

DOMHOFF, G.W. (1983). Who Rules America Now? A View for the '80s, Prentice-Hall, Englewood Cliffs, NJ.

DOMHOFF, G.W. (1980). Power Structure Research, Sage Publications, Beverly Hills, CA.

DOMHOFF, G.W. (1967). Who Rules America? Prentice-Hall, Englewood Cliffs, NJ.

DOUGHERTY, T.W.; TURBAN, D.B.; CALLENDER, J.C. (1994). Confirming First Impressions in the Employment Interview – A Field Study of Interviewer Behavior, *Journal of Applied Psychology*, 79, 659-665.

DREITZEL, H. P. (1962). Elitebegriff und Sozialstruktur – Eine soziologische Begriffsanalyse, Ferdinand Enke Verlag, Stuttgart.

DRYDAKIS, N.; VLASSIS, M. (2010). Ethnic Discrimination in the Greek Labour Market – Occupational Access, Insurance Coverage and Wage Offers, *The Manchester School*, 78, 201-218.

DUFFOURC, D. (2006). State-Funded College Scholarships – General Definitions and Characteristics, *Review of Policy Research*, 23, 235-248.

DUNN, W.S.; MOUNT, M.K.; BARRICK, M.R.; ONES, D.S. (1995). Relative Importance of Personality and General Mental Ability in Managers' Judgments of Applicant Qualifications, *Journal of Applied Psychology*, 80, 500-509.

DYE, T.R. (1994). Who's Running America? The Clinton Years, Prentice-Hall, Englewood Cliffs, NJ.

DYE, T.R. (1989). Who's Running America? The Bush Era, Prentice-Hall, Englewood Cliffs, NJ.

DYE, T.R. (1986). Who's Running America? The Conservative Years, Prentice-Hall, Englewood Cliffs, NJ.

DYE, T.R. (1983). Who's Running America? The Reagan Years, Prentice-Hall, Englewood Cliffs, NJ.

DYE, T.R. (1979). Who's Running America? The Carter Years, Prentice-Hall, Englewood Cliffs, NJ.

DYE, T.R. (1976). Who's Running America? Prentice-Hall, Englewood Cliffs, NJ.

EAGLY, A.H.; ASHMORE, R.D.; MAKHIJANI, M.G.; LONGO, L.A. (1991). What Is Beautiful Is Good, But… – A Meta-Analytic Review of Research on the Physical Attractiveness Stereotype, *Psychological Bulletin*, 110, 109-128.

EDWARDS, J.R. (1991). Person-Job fit – A Conceptual Integration, Literature Review and Methodological Critique, in Cooper, C.L.; Robertson, I.T. (Editors): International Review of Industrial and Organizational Psychology, Volume 6, John Wiley & Sons, Oxford, 283-357.

ELLIOTT, A.G.P. (1981). Sex and Decision Making in the Selection Interview – A Real-Life Study, *Journal of Occupational Psychology*, 54, 265-273.

ELLIS, A.P.J.; WEST, B.J.; RYAN, A.M.; DESHON, R.P. (2002). The Use of Impression Management Tactics in Structured Interviews – A Function of Question Type? *Journal of Applied Psychology*, 87, 1200-1208.

ENCYCLOPEDIA BRITANNICA (1975). The New Encyclopedia Britannica in 30 Volumes – Micropedia Volume III, 15th Ed., Encyclopedia Britannica Inc.

ESMAIL, A.; EVERINGTON, S. (1997). Asian Doctors Are Still Being Discriminated against, *British Medical Journal,* 314, 1619.

EVULEOCHA, S.U.; UGBAH, S.D.; LAW, S. (2009). Recruiter Perceptions of Information that Employment References Should Provide to Assist in Making Selection Decisions, *Journal of Employment Counseling,* 46, 98-106.

FAHR, R. (2006). Wage Effects of Social Norms – Evidence of Deviations from Peers' Body Mass in Europe, *IZA Discussion Paper* No. 2323.

FAMA, E.F. (1980). Agency Problems and the Theory of the Firm, *Journal of Political Economy*, 88, 288-307.

FAMA, E.F.; JENSEN, M.C. (1983). Agency Problems and Residual Claims, *Journal of Law and Economics*, 26, 327-349.

FARINA, A.; HAGELAUER, H.D. (1975). Sex and Mental Illness – The Generosity of Females, *Journal of Consulting and Clinical Psychology*, 43, 122.

FEDERAL STUDENT AID – AN OFFICE OF THE U.S. DEPARTMENT OF EDUCATION, accessed October 24, 2012 at: http://studentaid.ed.gov/about.

FEILD, H.S.; HOLLEY, W.H. (1976). Résumé Preparation – An Empirical Study of Personnel Managers' Perceptions, *The Vocational Guidance Quarterly,* 24, 229-237.

FELBER, W. (1986). Eliteforschung in der Bundesrepublik Deutschland – Analyse, Kritik, Alternativen, B.G. Teubner, Stuttgart.

FERGUSON, E.; JAMES, D.; O'HEHIR, F.; SANDERS, A. (2003). Pilot Study of the Roles of Personality, References, and Personal Statements in Relation to Performance over the Five Years of a Medical Degree, *British Medical Journal,* 326, 429-432.

FIEDLER, M.; WELPE, I.M. (2008). 'If You Don't Know What Port You Are Sailing to, No Wind Is Favorable' – Appointment Preferences of Management Professors, *Schmalenbach Business Review*, 60, 4-31.

FIGLIO, D.; HART, C.M.D.; METZGER, M. (2010). Who Uses a Means-Tested Scholarship, and What Do They Choose? *Economics of Education Review,* 29, 301-317.

FINKELSTEIN, L.M.; BURKE, M.J. (1998). Age Stereotyping at Work – The Role of Rater and Contextual Factors on Evaluations of Job Applicants, *The Journal of General Psychology*, 125, 317-345.

FINKELSTEIN, L.M.; BURKE, M.J.; RAJU, N.S. (1995). Age Discrimination in Simulated Employment Contexts – An Integrative Analysis, *Journal of Applied Psychology*, 80, 652-663.

FIRESTONE, R.; YANOFF, A.; MONTENEGRO, A. (2002). Employment Discrimination against Latinos in the Washington, D.C. Area, Arnold & Porter, Washingtion DC.

FIRTH, M. (1982). Sex Discrimination in Job Opportunities for Women, *Sex Roles,* 8, 891-901.

FIRTH, M. (1981). Racial Discrimination in the British Labor Market, *Industrial and Labor Relations Review*, 34, 265-272.

FISCHER, L. (2006). Studium – und darüber hinaus? Gesellschaftliches Engagement deutscher Studierender, HISBUS Kurzinformation Nr. 15, HIS: Projektbericht Oktober 2006.

FLOSSMANN, A.L.; PIATEK, R.; WICHERT, L. (2007). Going Beyond Returns to Education – The Role of Noncognitive Skills on Wages in Germany*,* Working Paper, University of Konstanz.

FORMAZIN, M.; SCHROEDERS, U.; KÖLLER, O.; WILHELM, O.; WESTMEYER, H. (2011). Studierendenauswahl im Fach Psychologie – Testentwicklung und Validitätsbefunde, *Psychologische Rundschau,* 62, 221-236.

FORSYTHE, S.M. (1990). Effects of Applicant's Clothing on Interviewer's Decision to Hire, *Journal of Applied Social Psychology*, 20, 1579-1595.

FREEMAN, R. B. (2005). Fellowship Stipend Support and the Supply of Science and Engineering Students – NSF Graduate Research Fellowships, *American Economic Review: Papers and Proceedings,* 95, 61-65.

FRANKE, N.; GRUBER, M; HARHOFF, D.; HENKEL, J. (2006). What You Are Is What You Like – Similarity Biases in Venture Capitalists' Evaluations of Start-Up Teams, *Journal of Business Venturing*, 21, 802-826.

FUSILIER, M.R.; HITT, M.A. (1983). Effects of Age, Race, Sex, and Employment Experience on Students' Perceptions of Job Applications, *Perceptual and Motor Skill,* 57, 1127-1134.

GARCÍA, M.F.; POSTHUMA, R.A.; COLELLA, A. (2008). Fit Perceptions in the Employment Interview – The Role of Similarity, Liking, and Expectations, *Journal of Occupational and Organizational Psychology*, 81, 173-189.

GATEWOOD, R.D.; FEILD, H.S. (1987). Human Resource Selection, The Dryden Press, Chicago.

GAUGLER, E.; MARTIN, C.; SCHNEIDER, B. (1995). Zusatzkriterien bei der Rekrutierung von Hochschulabsolventen, insbesondere Ehrenamtliche Tätigkeiten während des Studiums – Ergebnisse einer empirischen Studie, Mannheim, Forschungsstelle für Betriebswirtschaft und Sozialpraxis e.V. (FBS) Schriftenreihe, Band 51.

GIBSON, K.J.; ZERBE, W.J.; FRANKEN, R.E. (1993). The Influence of Rater and Ratee Age on Judgments of Work-Related Attributes, *The Journal of Psychology*, 127, 271-280.

GILMORE, D.C.; BEEHR, T.A.; LOVE, K.G. (1986). Effects of Applicant Sex, Applicant Physical Attractiveness, Type of Rater and Type of Job Interview Decisions, *Journal of Occupational Psychology*, 59, 103-109.

GILMORE, D.C.; FERRIS, G.R. (1989). The Effects of Applicant Impression Management Tactics on Interviewer Judgments, *Journal of Management*, 15, 557-564.

GINKEL, R.W.; DAVIS, S.E.; MICHAEL, P.G. (2010). An Examination of Inclusion and Exclusion Criteria in the Predoctoral Internship Selection Process, *Training and Education in Professional Psychology,* 4, 213-218.

GNEEZY, U.; LIST, J.A.; PRICE, M.K. (2012). Toward an Understanding of Why People Discriminate – Evidence from a Series of Natural Field Experiments, *NBER Working Paper* No. 17855.

GOLDBERG, C.B. (2005). Relational Demography and Similarity-Attraction in Interview Assessments and Subsequent Offer Decisions – Are We Missing Something? *Group & Organization Management*, 30, 597-624.

GOLDBERG, L.R. (1993). The Structure of Phenotypic Personality Traits, *American Psychologist*, 48, 26-34.

GOLDBERG, A.; MOURINHO, D.; KULKE, U. (1996). Labour Market Discrimination against Foreign Workers in Germany, *International Migration Papers* No. 7, International Labour Office, Geneva.

GOLDIN, C.; ROUSE, C. (2000). Orchestrating Impartiality – The Impact of "Blind" Auditions on Female Musicians, *American Economic Review,* 90, 715-741.

GOLDTHORPE, J. (1982). On the Service Class, Its Formation and Future, in Giddens, A.; Mackenzie, G. (Editors): Social Class and the Division of Labour – Essays in Honor of Ilya Neustadt, Cambridge University Press, Cambridge, 162-185.

GORDON, R.A. (2012). Applied Statistics for the Social and Health Sciences, Routledge, New York, NY.

GORDON, R.A.; ARVEY, R.D. (2004). Age Bias in Laboratory and Field Settings – A Meta-Analytic Investigation, *Journal of Applied Social Psychology*, 34, 468-492.

GRAVES, L.M; POWELL, G.N. (1996). Sex Similarity, Quality of the Employment Interview and Recruiters' Evaluation of Actual Applicants, *Journal of Occupational and Organizational Psychology*, 69, 243-261.

GRAVES, L.M; POWELL, G.N. (1995). The Effect of Sex Similarity on Recruiters' Evaluations of Actual Applicants – A Test of the Similarity-Attraction Paradigm, *Personnel Psychology,* 48, 85-98.

GRIFFETH, R.W.; BEDEIAN, A.G. (1989). Employee Performance Evaluations – Effects of Ratee Age, Rater Age, and Rater Gender, *Journal of Organizational Behavior*, 10, 83-90.

GROSSMAN, S.J.; HART, O.D. (1983). An Analysis of the Principal-Agent Problem, *Econometrica*, 51, 7-45.

GUJARATI, D.N.; PORTER, D.C. (2009). Basic Econometrics, 5th Ed., International Edition, McGrawHill, Boston.

HAIR, J.F. JR.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E. (2010). Multivariate Data Analysis – A Global Perspective, 7th Ed., Pearson Education, Upper Saddle River, NJ.

HAKEL, M.D.; DOBMEYER, T.W.; DUNNETTE, M.D. (1970). Relative Importance of Three Content Dimensions in Overall Suitability Ratings of Job Applicants' Résumés, *Journal of Applied Psychology,* 54, 65-71.

HAKEL, M.D.; OHNESORGE, J.P.; DUNNETTE, M.D. (1970). Interviewer Evaluations of Job Applicants' Résumés as a Function of the Qualifications of the Immediately Preceding Applicants – An Examination of Contrast Effects, *Journal of Applied Psychology*, 54, 27-30.

HALE, B. (2009). What's so Moral about the Moral Hazard? *Public Affairs Quarterly*, 23, 1-25.

HALL, J.A. (1978). Gender Effects in Decoding Nonverbal Cues, *Psychological Bulletin*, 85, 845-857.

HAMERMESH, D.S. (2011). Beauty Pays – Why Attractive People Are More Successful, Princeton University Press, Princeton, NJ.

HAMERMESH, D.S.; BIDDLE, J.E. (1994). Beauty and the Labor Market, *American Economic Review*, 84, 1174-1194.

HAMERMESH, D.S.; PARKER, A.M. (2003). Beauty in the Classroom – Professors' Pulchritude and Putative Pedagogical Productivity, *NBER Working Paper* No. 9853.

HAN, E.; NORTON, E.C.; STEARNS, S.C. (2009). Weight and Wages – Fat versus Lean Paychecks, *Health Economics*, 18, 535-548.

HARCOURT, J.; KRIZAN, A.C. (1989). A Comparison of Résumé Content Preferences of Fortune 500 Personnel Administrators and Business Communication Instructors, *The Journal of Business Communication,* 26, 177-190.

HAROLD, C.M.; MCFARLAND, L.A.; WEEKLEY, J.A. (2006). The Validity of Verifiable and Non-Verifiable Biodata Items – An Examination across Applicants and Incumbents, *International Journal of Selection and Assessment,* 14, 336-346.

HARPER, B. (2000). Beauty, Stature and the Labour Market – A British Cohort Study, *Oxford Bulletin of Economics and Statistics*, 62, Issue Supplement, 771-800.

HARRIS, M.B.; HARRIS, R.J.; BOCHNER, S. (1982). Fat, Four-Eyed, and Female – Stereotypes of Obesity, Glasses, and Gender, *Journal of Applied Social Psychology*, 12, 503-516.

HARTMANN, M. (2007). Eliten und Macht in Europa – Ein internationaler Vergleich, Campus Verlag, Frankfurt.

HARTMANN, M. (2004). Elitesoziologie – Eine Einführung, Campus Verlag, Frankfurt.

HARTMANN, M. (2002). Der Mythos von den Leistungseliten – Spitzenkarrieren und soziale Herkunft in Wirtschaft, Politik, Justiz und Wissenschaft, Campus Verlag, Frankfurt.

HARTMANN, M. (2001). Die Bildungsexpansion in der Bundesrepublik Deutschland und der Zugang zu Spitzenpositionen in Wirtschaft, Wissenschaft, Justiz und Politik – Zum Zusammenhang zwischen sozialer Herkunft, Promotion und dem Aufstieg in die politische, juristische, wissenschaftliche und wirtschaftliche Elite, *Deutsche Forschungsgemeinschaft (DFG)-Forschungsbericht*, Darmstadt.

HARVEY, L.; MOON, S.; GEALL, V.; BOWER, R. (1997). Graduate's Work – Organisational Change and Students' Attributes, *Centre for Research into Quality (CRQ)*, Online version accessed October 24, 2012 at: http://www0.bcu.ac.uk/crq/publications/gw/gwovf.html

HECKMAN, J.J. (1998). Detecting Discrimination, *Journal of Economic Perspectives,* 12, 101-116.

HECKMAN, J.J.; SIEGELMAN, P. (1993). The Urban Institute Audit Studies – Their Methods and Findings, in Fix, M.E.; Struyk, R.J. (Editors): Clear and Convincing Evidence – Measurement of Discrimination in America, The Urban Institute Press, Washington, DC, 187-258.

HEILMAN, M.E.; SARUWATARI, L.R. (1979). When Beauty is Beastly – The Effects of Applicant Appearance and Sex on Evaluations of Job Applicants for Managerial and Nonmanagerial Jobs, *Organizational Behavior and Human Performance*, 23, 360-372.

HELLSTRÖM, J. (2011). Conditional Hypotheses in Comparative Social Science – Mixed-Method Approaches to Middle-Sized Data Analysis, *Methodological Innovations Online*, 6, 71-102.

HENEMANN III, H.G.; SCHWAB, D.P.; HUETT, D.L.; FORD, J.J. (1975). Interviewer Validity as a Function of Interview Structure, Biographical Data, and Interviewee Order, *Journal of Applied Psychology*, 60, 748-753.

HENSS, R. (1992). 'Spieglein, Spieglein an der Wand…' Geschlecht, Alter und physische Attraktivität, Psychologie Verlags Union, Weinheim.

HEUBLEIN, U.; HUTZSCH, C.; SCHREIBER, J.; SOMMER, D. (2011). Internationale Mobilität im Studium 2009 – Ergebnisse einer Wiederholungsbefragung zu studienbezogenen Aufenthalten deutscher Studierender in anderen Ländern, HIS: Projektbericht September 2011.

HEWSTONE, M.; RUBIN, M.; WILLIS, H. (2002). Intergroup Bias, *Annual Review of Psychology,* 53, 575-604.

HIGGINS, C.A.; JUDGE, T.A. (2004). The Effect of Applicant Influence Tactics on Recruiter Perceptions of Fit and Hiring Recommendations – A Field Study, *Journal of Applied Psychology*, 89, 622-632.

HINKS, T. (2002). Gender Wage Differentials and Discrimination in the New South Africa, *Applied Economics,* 34, 2043-2052.

HIS/HOCHSCHULINFORMATIONSSYSTEM (2012). Durchschnittliche Schulabschlussnote der Studienberechtigten 2006, 2008 und 2010 nach Land des Erwerbs der Hochschulreife – HIS-Studienberechtigtenbefragung 2006-2010, Hannover.

HOFFMANN-LANGE, U. (2007). Methods of Elite Research. in Dalton, R.J.; Klingemann, H.-D. (Editors): Oxford Handbook of Political Behavior, Oxford University Press, Oxford, 910-927.

HOFFMANN-LANGE, U. (1987). Surveying National Elites in the Federal Republic of Germany, in Moyser, G.; Wagstaffe, M. (Editors): Research Methods for Elite Studies, Allen & Unwin, London, 27-47.

HOFFMANN-LANGE, U. (1992). Eliten, Macht und Konflikt in der Bundesrepublik Deutschland, Leske + Budrich, Opladen.

HOLLAND, J.L. (1985). Making Vocational Choices – A Theory of Vocational Personalities and Work Environments, 2nd Ed., Prentice-Hall, Englewood Cliffs, NJ.

HOMSTRÖM, B. (1982). Moral Hazard in Teams, *The Bell Journal of Economics*, 13, 324-340.

HOLMSTRÖM, B. (1979). Moral Hazard and Observability, *The Bell Journal of Economics*, 10, 74-91.

HOMBURG, C.; SCHNEIDER, J.; FASSNACHT, M. (2002). Opposites Attract, but Similarity Works – A Study of Interorganizational Similarity in Marketing Channels, *Institute for the Study of Business Marktes (ISBM) Report* 6-2002.

HOSMER, D.W.; LEMESHOW, S. (2000). Applied Logistic Regression, 2nd Ed., John Wiley & Sons, New York, NY.

HOSODA, M.; STONE-ROMERO, E.F.; COATS, G. (2003). The Effects of Physical Attractiveness on Job-Related Outcomes – A Meta-Analysis of Experimental Studies, *Personnel Psychology*, 56, 431-462.

HOWARD, J.L.; FERRIS, G.R. (1996). The Employment Interview Context – Social and Situational Influences on Interviewer Decisions, *Journal of Applied Social Psychology*, 26, 112-136.

HUBBUCK, J.; CARTER, S. (1980). Half a Chance? A Report on Job Discrimination against Young Blacks in Nottingham, Commission for Racial Equality, London.

HUFFCUTT, A.I.; VAN IDDEDKINGE, C.H.; ROTH, P.L. (2011). Understanding Applicant Behavior in Employment Interviews – A Theoretical Model of Interviewee Performance, *Human Resource Management Review*, 21, 353-367.

HUNTER, J.E.; HUNTER, R.F. (1984). Validity and Utility of Alternative Predictors of Job Performance, *Psychological Bulletin,* 96, 72-98.

HURTZ, G.M.; DONOVAN, J.J. (2000). Personality and Job Performance – The Big Five Revisited, *Journal of Applied Psychology*, 85, 869-879.

HUTCHINSON, K.L. (1984). Personnel Administrators' Preferences for Résumé Content – A Survey and Review of Empirically Based Conclusions, *The Journal of Business Communication,* 21, 5-14.

HUTCHINSON, K.L.; BREFKA, D.S. (1997). Personnel Administrators' Preferences for Résumé Content – Ten Years after, *Business Communication Quarterly,* 60, 67-75.

INVESTOPEDIA (2012). Definition of 'Stipend', accessed October 24, 2012 at: http://www.investopedia.com/terms/s/stipend.asp#axzz24N6kk3OL.

ISSERSTEDT, W.; MIDDENDORFF, E.; FABIAN, G.; WOLTER, A. (2007). Die wirtschaftliche und soziale Lage der Studierenden in der Bundesrepublik Deutschland 2006 – 18. Sozialerhebung des Deutschen Studentenwerks durchgeführt durch HIS Hochschul-Informations-System, Bundesministerium für Bildung und Forschung (BMBF), Bonn.

JACKSON, S.F.; BRETT, J.F.; SESSA, V.I.; COOPER, D.M.; JULIN, J.A.; PEYRONNIN, K. (1991). Some Differences Make a Difference – Individual Dissimilarity and Group Heterogeneity as Correlates of Recruitment, Promotion, and Turnover, *Journal of Applied Psychology*, 76, 675-689.

JACKSON, L.A.; HUNTER, J.E.; HODGE, C.N. (1995). Physical Attractiveness and Intellectual Competence – A Meta-Analytic Review, *Social Psychology Quarterly*, 58, 108-122.

JANSEN, K.J. ; KRISTOF-BROWN, A. (2006). Toward a Multidimensional Theory of Person-Environment Fit*, Journal of Managerial Issues*, 18, 193-212.

JENSEN, M.C.; MECKLING, W.H. (1976). Theory of the Firm – Managerial Behavior, Agency Costs and Ownership Structure, *Journal of Financial Economics*, 3, 305-360.

JOHNSON, R.W.; NEUMARK, D. (1997). Age Discrimination, Job Separations, and Employment Status of Older Workers – Evidence from Self-Reports, *Journal of Human Resources*, 32, 779-811.

JOWELL, R.; PRESCOTT-CLARKE, P. (1970). Racial Discrimination and White-Collar Workers in Britain, *Race Class*, 11, 397-417.

JUDGE, T.A.; HIGGINS, C.A. (1998). Affective Disposition and the Letter of Reference, *Organizational Behavior and Human Decision Processes,* 75, 207-221.

JUDGE, T.A.; HIGGINS, C.A.; CABLE, D.M. (2000). The Employment Interview – A Review of Recent Research and Recommendations for Future Research, *Human Resource Mangagement Review*, 10, 383-406.

KAAS, L.; MANGER, C. (2010). Ethnic Discrimination in Germany's Labour Market – A Field Experiment*, IZA Discussion Paper* No. 4741.

KACMAR, K.M.; DELERY, J.E.; FERRIS, G.R. (1992). Differential Effectiveness of Applicant Impression Management Tactics on Employment Interview Decisions, *Journal of Applied Social Psychology*, 22, 1250-1272.

KAINA, V. (2009). Eliteforschung, in Kaina, V.; Römmele A. (Editors): Politische Soziologie – Ein Studienbuch, VS Verlag für Sozialwissenschaften, Wiesbaden, 385-419.

KAINA, V. (2006). Was sind Eliten? in Gabriel, O.W.; Neuss, B.; Rüther, G. (Editors): Eliten in Deutschland – Bedeutung – Macht – Verantwortung, Bundeszentrale für Politische Bildung, Bonn, 41-56.

KELLER, S. (1963). Beyond the Ruling Class – Strategic Elites in Modern Society, Random House, New York, NY.

KITE, M.E.; STOCKDALE, G.D.; WHITLEY JR., B.E.; JOHNSON, B.T. (2005). Attitudes Toward Younger and Older Adults – Updated Meta-Analytic Review, *Journal of Social Issues*, 61, 241-266.

KLEIN, M.; ROSAR, U. (2005). Physische Attraktivität und Wahlerfolg – Eine empirische Analyse am Beispiel der Wahlkreiskandidaten bei der Bundestagswahl 2002, *Politische Vierteljahresschrift*, 46, 263-287.

KLEINMANN, M.; KLEHE, U.-C. (2010). Selling Oneself – Construct and Criterion-Related Validity of Impression Management in Structured Interviews, *Human Performance*, 24, 29-46.

KMK/SEKRETARIAT DER STÄNDIGEN KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND (2006). Abiturnoten an allgemein bildenden Gymnasien und integrierten Gesamtschulen – Schuljahr 2004/2005, Bonn.

KNOUSE, S.B. (1994a). Impressions of the Résumé – The Effects of Applicant Education, Experience and Impression Management, *Journal of Business and Psychology*, 9, 33-45.

KNOUSE, S.B. (1994b). Improving the Letter of Recommendation, *Journal of Employment Counseling,* 31, 105-109.

KNOUSE, S.B. (1983). The Letter of Recommendation – Specificity and Favorability of Information, *Personnel Psychology,* 36, 331-341.

KNOUSE, S.B.; TANNER, J.R.; HARRIS, E.W. (1999). The Relation of College Internships, College Performance, and Subsequent Job Opportunity, *Journal of Employment Counseling,* 36, 35-43.

KOEDEL, C.; TYHURST, E. (2012). Math Skills and Labor-Market Outcomes – Evidence from a Résumé-Based Field Experiment, *Economics of Education Review,* 31, 131-140.

KOHN, A.R.; FIEDLER, F.E. (1961). Age and Sex Differences in the Perception of Persons, *Sociometry*, 24, 157-164.

KOLLE, A. (2012a). Ethnic Discrimination in the Recruitment Process – A Field Experiment in the German Labor Market, Unpublished Working Paper.

KOLLE, A. (2012b). Gender Discrimination in the Recruitment Process – A Field Experiment in the German Labor Market, *Colloquium on Personnel Economics* Working Paper.

KRAUSE, A.; RINNE, U.; ZIMMERMANN, K.F. (2010). Anonymisierte Bewerbungsverfahren, *IZA Research Report* No. 27.

KRISTOF, A. L. (1996). Person-Organization Fit – An Integrative Review of Its Conceptualizations, Measurement, and Implications, *Personnel Psychology*, 49, 1-49.

KRISTOF-BROWN, A. L. (2000). Perceived Applicant Fit – Distinguishing between Recruiters' Perceptions of Person-Job and Person-Organization Fit, *Personnel Psychology*, 53, 643-671.

KRISTOF-BROWN, A.; BARRICK, M.R.; FRANKE, M. (2002). Applicant Impression Management – Dispositional Influences and Consequences for Recruiter Perceptions of Fit and Similarity, *Journal of Management*, 28, 27-46.

KRISTOF-BROWN, A.L.; ZIMMERMAN, R.D.; JOHNSON, E.C. (2005). Consequences of Individuals' Fit at Work – A Meta-Analysis of Person-Job, Person-Organization, Person-Group, and Person-Supervisor Fit, *Personnel Psychology*, 58, 281-342.

KUHN, P.; WEINBERGER, C. (2005). Leadership Skills and Wages, *Journal of Labor Economics*, 23, 395-436.

KWANTES, C.T.; LIN, I.Y.; GIDAK, N.; SCHMIDT, K. (2011). The Effect of Attire on Expected Occupational Outcomes for Male Employees, *Psychology of Men & Masculinity*, 12, 166-180.

LAHEY, J.N. (2008). Age, Women, and Hiring – An Experimental Study, *Journal of Human Resources*, 43, 30-56.

LANDRY, C.; LANGE, A.; LIST, J.A.; PRICE, M.K.; RUPP, N.G. (2005). Toward an Understanding of the Economics of Charity – Evidence from a Field Experiment, *NBER Working Paper* No. 11611.

LANDY, F.J.; BATES, F. (1973). Another Look at Contrast Effects in the Employment Interview, *Journal of Applied Psychology*, 58, 141-144.

LANGLOIS, J.H.; KALAKANIS, L.; RUBENSTEIN, A.J.; LARSON, A.; HALLAM, M.; SMOOT, M. (2000). Maxims or Myths of Beauty? A Meta-Analytic and Theoretical Review, *Psychological Bulletin*, 126, 390-423.

LASWELL, H.D. (1936). Politics – Who Gets What, When and How, McGraw-Hill, New York, NY.

LASWELL, H.D. (1934). World Politics and Personal Insecurity, University of Chicago Press, Chicago, IL.

LATHAM, G.P.; WEXLEY, K.N.; PURSELL, E.D. (1975). Training Managers to Minimize Rating Errors in the Observation of Behavior, *Journal of Applied Psychology*, 60, 550-555.

LAYTON, B.D.; INSKO, C.A. (1974). Anticipated Interaction and the Similarity-Attraction Effect, *Sociometry*, 37, 149-162.

LE BON, G. (2001). The Crowd – A Study of the Popular Mind, Batoche Books Kitchener, accessed October 24, 2012 at: http://socserv.mcmaster.ca/econ/ugcm/3ll3/lebon/Crowds.pdf

LEFKOWITZ, J. (2000). The Role of Interpersonal Affective Regard in Supervisory Performance Ratings – A Literature Review and Proposed Causal Model, *Journal of Occupational and Organizational Psychology*, 73, 67-85.

LEVASHINA, J.; CAMPION, M.A. (2006). A Model of Faking Likelihood in the Employment Interview, *International Journal of Selection and Management*, 14, 299-316.

LEVINSON, R.M. (1975). Sex Discrimination and Employment Practices – An Experiment with Unconventional Job Inquiries, *Social Problems,* 22, 533-543.

LEWIS, T.D.; SHIMERDA, T.A.; GRAHAM, G. (1983). What the Academic Advisor Needs to Know about Job Placement, *Journal of Accounting Education,* 1, 135-142.

LICHTENTHAL, J.D.; TELLEFSEN, T. (2001). Toward a Theory of Business Buyer-Seller Similarity, *Journal of Personal Selling & Sales Management*, 21, 1-14.

LIEN, D. (2007). The Role of Scholarships in Study Abroad Programs, *Education Economics,* 15, 203-213.

LIEN, D.; LIU, G. (2010). Financial Assistance for Study Abroad Students – An Economic Analysis, *International Review of Economics and Finance,* 19, 515-522.

LIEN, D.; WANG, Y. (2010). Optimal Design for Study-Abroad Scholarship – The Effect of Payback Policy, *Education Economics,* 18, 191-205.

LIN, T.-R.; DOBBINS, G.H.; FARH, J.-L. (1992). A Field Study of Race and Age Similarity Effects on Interview Ratings in Conventional and Situational Interviews, *Journal of Applied Psychology*, 77, 363-371.

LIST, J.A. (2004). The Nature and Extent of Discrimination in the Marketplace – Evidence from the Field, *Quarterly Journal of Economics,* 119, 49-89.

LODDER, L.; MCFARLAND, S.; WHITE, D. (2003). Racial Preference and Suburban Employment Opportunities – A Report on 'Matched-Pair' Tests of Chicago Area Retailers, Legal Assistance Foundation of Metropolitan Chicago and Chicago Urban League.

LOHER, B.T.; HAZER, J.T.; TSAI, A.; TILTON, K.; JAMES, J. (1997). Letters of Reference – A Process Approach, *Journal of Business and Psychology*, 11, 339-355.

LONDON, M.; POPLAWSKI, J.R. (1976). Effects of Information on Stereotype Development in Performance Appraisal and Interview Contexts, *Journal of Applied Psychology*, 61, 199-205.

LONG, J.S.; FREESE, J. (2006). Regression Models for Categorical Dependent Variables Using Stata, 2nd Ed., Stata Press, College Station, TX.

LÓPEZ BÓO, F.; ROSSI, M.A.; URZÚA, S. (2012). The Labor Market Return to an Attractive Face – Evidence from a Field Experiment, *IZA Discussion Paper* No. 6356.

LUXEN, M.F.; VAN DE VIIJVER, F.J.R. (2006). Facial Attractiveness, Sexual Selection, and Personnel Selection – When Evolved Preferences Matter, *Journal of Organizational Behavior*, 27, 241-255.

LYNN, M. (2009). Determinants and Consequences of Female Attractiveness and Sexiness – Realistic Tests with Restaurant Waitresses, *Archives of Sexual Bahavior*, 38, 737-745.

MACAN, T. (2009). The Employment Interview – A Review of Current Studies and Directions for Future Research, *Human Resource Management Review*, 19, 203-218.

MACHATZKE, J. (1997). Die Potsdamer Elitestudie – Positionsauswahl und Ausschöpfung, in Bürklin, W.; Rebenstorf, H. et al. (Editors): Eliten in Deutschland – Rekrutierung und Integration, Leske + Budrich, Opladen, 35-68.

MACK, D.; RAINEY, D. (1990). Female Applicants' Grooming and Personnel Selection, *Journal of Social Behavior and Personality*, 5, 399-407.

MANNHEIM, K. (1967). Mensch und Gesellschaft im Zeitalter des Umbaus, 2nd Ed., Verlag Gehlen, Bad Homburg.

MANSKI, C.F.; WISE, D.A. (1983). College Choice in America, Harvard University Press, Cambridge.

MANZ, W.; LUECK, H.E. (1968). Influence of Wearing Glasses on Personality Ratings – Crosscultural Validation of an Old Experiment, *Perceptual and Motor Skills*, 27, 704.

MARLOWE, C.M.; SCHNEIDER, S.L.; NELSON, C.E. (1996). Gender and Attractiveness Biases in Hiring Decisions – Are More Experienced Managers Less Biased? *Journal of Applied Psychology*, 81, 11-21.

MCCARTHY, J.M.; VAN IDDEKINGE, C.H.; CAMPION, M.A. (2010). Are Highly Structured Job Interviews Resistant to Demographic Similarity Effects? *Personnel Psychology,* 63, 325-359.

MCDANIEL, M.A.; SCHMIDT, F.L.; HUNTER, J.E. (1988). Job Experience Correlates of Job Performance, *Journal of Applied Psychology, 73*, 327-330.

MCDONALD, L.; CHEN, M.Y.T. (1994). The Youth Freeze and the Retirement Bulge – Older Workers and the Impending Labour Shortage, in Marshall, V.W.; McPherson, B. (Editors): Aging – Canadian Perspectives, Broadview Press/Journal of Canadian Studies, Peterborough, ON, 111-139.

MCDONALD, A.S.; NEWTON, P.E.; WHETTON, C.; HIGGS, S.E. (2001). A Pilot of Aptitude Testing for University Entrance, *NFER Draft Report*.

MCFADDEN, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior, in Zarembka, P. (Editor): Frontiers of Econometrics, Academic Press, New York, NY, 105-142.

MCFARLAND, L.A.; RYAN, A.M.; KRISKA, S.D. (2003). Impression Management Use and Effectiveness across Assessment Methods, *Journal of Management*, 29, 641-661.

MCFARLAND, L.A.; RYAN, A.M.; KRISKA, S.D. (2002). Field Study Investigation of Applicant Use of Influence Tactics in a Selection Interview, *The Journal of Psychology*, 136, 383-398.

MCFARLAND, L.A.; RYAN, A.M.; SACCO, J.M.; KRISKA, S.D. (2004). Examination of Structured Interview Ratings across Time – The Effects of Applicant Race, Rater Race, and Panel Composition, *Journal of Management,* 30, 435-452.

MCINTOSH, N; SMITH, D. J. (1974). The Extent of Racial Discrimination, *Political and Economic Planning Broadsheet* No. 547, PEP, London.

MCKELVEY, R.D.; ZAVOINA, W. (1975). A Statistical Model for the Analysis of Ordinal Level Dependent Variables, *The Journal of Mathematical Sociology*, 4, 103-120.

MCKINNEY, A.P.; CARLSON, K.D.; MECHAM III, R.L.; D'ANGELO, N.C., CONNERLEY, M.L. (2003). Recruiters' Use of GPA in Initial Screening Decisions – Higher GPAs Don't Always Make the Cut, *Personnel Psychology,* 56, 823-845.

MEAD, J.F. (1965). A Description of Union Scholarship Programs, *Monthly Labor Review,* 88, 508-516.

MENARD, S. (2002). Applied Logistic Regression Analysis, 2nd Ed., Sage Publications, Thousand Oaks, Series: Quantitative Applications in the Social Sciences No. 07-106.

MERKER, L. (2009). Engagement, Studienerfolg und Berufserfolg – Eine Untersuchung über Zusammenhänge im Lebenslauf von Absolventen der Betriebswirtschaftslehre an der Universität Bayreuth, Dissertation.

MERKER, L.; KÜHLMANN, T.M. (2010). Was wirklich zählt, *Personal,* Heft 02/2010, 39-41.

MERTON, R.K. (1968). The Matthew Effect in Science, *Science*, 159, 59-63.

MEYERS GROSSES TASCHENLEXIKON (2003). Meyers Grosses Taschenlexikon in 26 Bänden – Band 6 Eih-Fau, 9th Ed., Bibliographisches Institut & F.A. Brockhaus, Mannheim.

MICHELS, R. (1970). Zur Soziologie des Parteiwesens in der modernen Demokratie – Untersuchungen über die oligarchischen Tendenzen des Gruppenlebens, 2nd Ed., Alfred Kröner Verlag, Stuttgart.

MIDDENDORFF, E.; ISSERSTEDT, W.; KANDULLA, M. (2009). Das soziale Profil in der Begabtenförderung – Ergebnisse einer Online-Befragung unter allen Geförderten der elf Begabtenförderungswerke im Oktober 2008, HIS-Projektbericht.

MILLS, C. W. (1958). The Structure of Power in American Society, *The British Journal of Sociology,* 9, 29-41.

MILLS, C. W. (1956). The Power Elite, Oxford University Press, New York, NY.

MOBIUS, M.M; ROSENBLAT, T.S. (2006). Why Beauty Matters, *American Economic Review*, 96, 222-235.

MONKS, J. (2009). The Impact of Merit-Based Financial Aid on College Enrollment – A Field Experiment, *Economics of Education Review*, 28, 99-106.

MONTOYA, R.M.; HORTON, R.S.; KIRCHNER, J. (2008). Is Actual Similarity Necessary for Attraction? A Meta-Analysis of Actual and Perceived Similarity, *Journal of Social and Personal Relationships*, 25, 889-922.

MOSCA, G. (1939). The Ruling Class (*Elementi die Scienzia Politica*), Translated, Edited and Revised Edition, McGraw-Hill, New York and London.

MOSEL, J.N.; GOHEEN, H.W. (1959). The Employment Recommendation Questionnaire – III. Validity of Different Types of References, *Personnel Psychology,* 12, 469-478.

MOSER, K.; RHYSSEN, D. (2001). Referenzen als eignungsdiagnostische Methode, *Zeitschrift für Arbeits- und Organisationspsychologie,* 45, 40-46.

MOUNT, M.K.; BARRICK, M.R.; STEWART, G.L. (1998). Five-Factor Model of Personality and Performance in Jobs Involving Interpersonal Interactions, *Human Performance*, 11, 145-165.

MUCHINSKY, P.M. (1979). The Use of Reference Reports in Personnel Selection – A Review and Evaluation, *Journal of Occupational Psychology,* 52, 287-297.

MUCHINSKY, P.M.; HARRIS, S.L. (1977). The Effect of Applicant Sex and Scholastic Standing on the Evaluation of Job Applicant Résumés in Sex-Typed Occupations, *Journal of Vocational Behavior*, 11, 95-108.

MUELLER, G.; PLUG, E.J.S. (2006). Estimating the Effect of Personality on Male and Female Earnings, *Industrial and Labor Relations Review*, 60, 3-22.

NEMANICK, R.C. JR.; CLARK, E.M. (2002). The Differential Effects of Extracurricular Activities on Attributions in Résumé Evaluation, *International Journal of Selection and Assessment*, 10, 206-217.

NEUMARK, D. (2011). Detecting Discrimination in Audit and Correspondence Studies, UCI Center for Economics & Public Policy and Department of Economics, NBER, and IZA, Working Paper August 2011.

NEUMARK, D. WITH THE ASSISTANCE OF BANK, R.J.; VAN NORT, K.D.(1996). Sex Discrimination in Restaurant Hiring – An Audit Study, *Quarterly Journal of Economics,* 111, 915-941.

NICKLIN, J.M.; ROCH, S.G. (2008). Biases Influencing Recommendation Letter Contents – Phsyical Attractiveness and Gender, *Journal of Applied Social Psychology,* 38, 3053-3074.

NORTON, E.C.; WANG, H.; AI, C. (2004). Computing Interaction Effects and Standard Errors in Logit and Probit Models, *The Stata Journal*, 4, 154-167.

NUNES, A.; SELIGMAN, B. (2000). A Study of the Treatment of Female and Male Applicants by San Francisco Bay Area Auto Service Shops, Discrimination Research Center of the Impact Fund, Berkeley, CA.

NUNES, A.; SELIGMAN, B. (1999). Treatment of Caucasian and African-American Applicants by San Francisco Bay Area Employment Agencies – Results of a Study Utilizing 'Testers', Discrimination Research Center of the Impact Fund, Berkeley, CA.

NYHUS, E.K.; PONS, E. (2005). The Effects of Personality on Earnings, *Journal of Economic Psychology*, 26, 363-384.

OAXACA, R. (1973). Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review,* 14, 693-709.

OECD/ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (1998). Workforce Ageing – Consequences and Policy Responses, *Ageing Working Papers*, 4.1, Organisation for Economic Co-operation and Development (OECD).

OLIAN, J.D.; SCHWAB, D.P.; HABERFELD, Y. (1988). The Impact of Applicant Gender Compared to Qualifications on Hiring Recommendations – A Meta-Analysis of Experimental Studies, *Organizational Behavior and Human Decision Processes,* 41, 180-195.

OPHEIM, V. (2006). Changing the System of Student Support in Norway – Were Policy Goals Met? *Journal of Higher Education Policy and Management,* 28, 277-287.

OPITZ, C. (2005). Hochschulen als Filter für Humankapital – Bildung und Karriere von High Potentials in den USA, Frankreich und Deutschland, Deutscher UNiversitäts-Verlag, Wiesbaden.

O'REILLY III, C.A.; CHATMAN, J.; CALDWELL, D.F. (1991). People and Organizational Culture – A Profile Comparison Approach to Assessing Person-Organization Fit, *The Academy of Management Journal*, 34, 487-516.

PAGER, D. (2007). The Use of Field Experiments for Studies of Employment Discrimination – Contributions, Critiques, and Directions for the Future, *The ANNALS of the American Academy of Political and Social Science*, 609, 104-133.

PAGER, D. (2003). The Mark of a Criminal Record, *American Journal of Sociology,* 108, 937-975.

PAGER, D.; WESTERN, B. (2005). Discrimination in Low-Wage Labor Markets – Results from an Experimental Audit Study in New York City, Submission to the Population Association of America Annual Meetings.

PAGER, D.; WESTERN, B.; BONIKOWSKI, B. (2009). Discrimination in a Low-Wage Labor Market – A Field-Experiment, *IZA Discussion Paper* No. 4469.

PALLETT, P.M.; LINK, S.; LEE, K. (2010). New 'Golden' Ratios for Facial Beauty, *Vision Research*, 50, 149-154.

PAMPEL, F.C. (2000). Logistic Regression – A Primer, Sage Publications, Thousand Oaks, Series: Quantitative Applications in the Social Sciences No. 07-132.

PARETO, V. (1935). The Mind and Society (*Trattato di Sociologica generale*), Volume 3: Theory of Derivations, 1120-1432, Harcourt Brace Jovanovich, New York, NY, accessed October 24, 2012 at: http://solomon.soth.alexanderstreet.com/cgi-bin/asp/philo/soth/getdoc.pl?S10023923-D000004

PARRETT, M.B. (2003). Beauty and the Labor Market – Evidence from a Tipping Data Set, in Parrett, M.B. (Editor): The Give and Take on Restaurant Tipping, Dissertation, Blacksburg, VA, 34-51, accessed October 24, 2012 at: http://scholar.lib.vt.edu/theses/available/etd-10172003-021146/unrestricted/ETDDissertation.pdf

PARSONS, C.K.; LIDEN, R.C. (1984). Interviewer Perceptions of Applicant Qualifications – A Multivariate Field Study of Demographic Characteristics and Nonverbal Cues, *Journal of Applied Psychology*, 69, 557-568.

PAUNONEN, S.V.; JACKSON, D.N.; OBERMAN, S.M. (1987). Personnel Selection Decisions – Effects of Applicant Personality and the Letter of Reference, *Organizational Behavior and Human Decision Processes*, 40, 96-114.

PEETERS, H.; LIEVENS, F. (2006). Verbal and Nonverbal Impression Management Tactics in Behavior Description and Situational Interviews, *International Journal of Selection and Assessment*, 14, 206-222.

PERES, S.H.; GARCIA, J.R. (1962). Validity and Dimensions of Descriptive Adjectives Used in Reference Letters for Engineering Applicants, *Personnel Psychology,* 15, 279-286.

PERRY, E.L.; BOURHIS, A.C. (1998). A Closer Look at the Role of Applicant Age in Selection Decisions, *Journal of Applied Social Psychology*, 28, 1670-1697.

PETERSON, K. (2008). Financial Aid Glossary, accessed October 24, 2012 at: http://www.fastweb.com/financial-aid/articles/815-financial-aid-glossary?page=5.

PETIT, P. (2007). The Effects of Age and Family Constraints on Gender Hiring Discrimination – A Field Experiment in the French Financial Sector, *Labour Economics,* 14, 371-391.

PHELPS, E.S. (1972). The Statistical Theory of Racism and Sexism, *American Economic Review,* 62, 659-661.

PHILLIPS, A.P.; DIPBOYE, R.L. (1989). Correlational Tests of Predictions from a Process Model of the Interview, *Journal of Applied Psychology*, 74, 41-52.

PIBAL, D.C. (1985). What Do You Look For in a Résumé? Criteria for Effective Résumés as Perceived by Personnel Directors, *Personnel Administrator,* 30, 119-123.

PICOT, A.; DIETL, H.; FRANCK, E.; FIEDLER, M.; ROYER, S. (2012). Organisation – Theorie und Praxis aus ökonomischer Sicht, 6th Ed., Schäffer-Poeschel, Stuttgart.

PINGITORE, R.; DUGONI, B.L.; TINDALE, R.S.; SPRING, B. (1994). Bias against Overweight Job Applicants in a Simulated Employment Interview, *Journal of Applied Psychology*, 79, 909-917.

PODOLNY, J.M. (1994). Market Uncertainty and the Social Character of Economic Exchange, *Administrative Science Quarterly*, 39, 458-483.

PORCU, M.; PUGGIONI, G.; TEDESCO, N. (2007). The Determinants of Graduates' Placement – Analysis of Interactions Using Boolean Logit Models, in Fabbris, L. (Editor): Effectiveness of University Education in Italy – Employability, Competences, Human Capital. Physica-Verlag, Heidelberg, 195-206.

POSTHUMA, R.A.; MORGESON, F.P.; CAMPION, M.A. (2002). Beyond Employment Interview Validity – A Comprehensive Narrative Review of Recent Research and Trends over Time, *Personnel Psychology,* 55, 1-81.

POWELL, G.N.; BUTTERFIELD, D.A. (2002). Exploring the Influence of Decision Makers' Race and Gender on Actual Promotions to Top Management, *Personnel Psychology*, 55, 397-428.

PREWETT-LIVINGSTON, A.J.; FEILD, H.S.; VERES III, J.G.; LEWIS, P.M. (1996). Effects of Race on Interview Ratings in a Situational Panel Interview, *Journal of Applied Psychology*, 81, 178-186.

PRICE, J.; WOLFERS, J. (2010). Racial Discrimination among NBA Referees, *Quarterly Journal of Economics*, 125, 1859-1887.

PROOST, K.; SCHREURS, B.; DE WITTE, K.; DEROUS, E. (2010). Ingratiation and Self-Promotion in the Selection Interview – The Effects of Using Single Tactics or a Combination of Tactics on Interviewer Judgments, *Journal of Applied Social Psychology*, 40, 2155-2169.

PURCELL, K.; WILTON, N.; ELIAS, P. (2003). Older and Wiser? Age and Experience in the Graduate Labour Market – Researching Graduate Careers Seven Years on, Research Paper No. 2.

PUTNAM, R.D. (1976). The Comparative Study of Political Elites, Prentice-Hall, Englewood Cliffs.

RAGIN, C.C. (1987). The Comparative Method – Moving beyond Qualitative and Quantitative Strategies, University of California Press, Berkely, CA.

RAND, T.M.; WEXLEY, K.N. (1975). Demonstration of the Effect 'Similar to Me' in Simulated Employment Interviews, *Psychological Reports*, 36, 535-544.

RAZA, S.M.; CARPENTER, B.N. (1987). A Model of Hiring Decisions in Real Employment Interviews, *Journal of Applied Psychology*, 72, 596-603.

REBENSTORF, H. (1997). Integration und Segmentation der Führungsschicht – Stratifikationstheoretische Determinanten, in Bürklin, W.; Rebenstorf, H. et al. (Editors): Eliten in Deutschland – Rekrutierung und Integration, Leske + Budrich, Opladen, 123-155.

REILLY, R.R.; CHAO, G.T. (1982). Validity and Fairness of Some Alternative Employee Selection Procedures, *Personnel Psychology,* 35, 1-62.

RENWICK, P.A.; TOSI, H. (1978). The Effects of Sex, Marital Status, and Education Background on Selection Decisions, *Academy of Management Journal,* 21, 93-103.

RIACH, P.A.; RICH, J. (2007a). An Experimental Investigation of Age Discrimination in the Spanish Labour Market, *IZA Discussion Paper* No. 2654.

RIACH, P.A.; RICH, J. (2007b). An Experimental Investigation of Age Discrimination in the English Labor Market, *IZA Discussion Paper* No. 3029.

RIACH, P.A.; RICH, J. (2006a). An Experimental Investigation of Sexual Discrimination in Hiring in the English Labor Market, *Advances in Economic Analysis & Policy,* 6, 1-20.

RIACH, P.A.; RICH, J. (2006b). An Experimental Investigation of Age Discrimination in the French Labour Market, *IZA Discussion Paper* No. 2522.

RIACH, P.A.; RICH, J. (2004). Deceptive Field Experiments of Discrimination – Are They Ethical? *KYKLOS, 57*, 457-470.

RIACH, P.A.; RICH, J. (2002). Field Experiments of Discrimination in the Market Place, *The Economic Journal,* 112, F480-F518.

RIACH, P.A.; RICH, J. (1991). Testing for Racial Discrimination in the Labour Market, *Cambridge Journal of Economics*, 15, 239-256.

RIACH, P.A.; RICH, J. (1987). Testing for Sexual Discrimination in the Labour Market, *Australian Economic Papers,* 26, 165-178.

RIGOL, G. W. (2003). Admission Decision-Making Models – How U.S. Institutions of Higher Education Select Undergraduate Students, The College Board.

ROEBKEN, H. (2010). Similarity Attracts – An Analysis of Recruitment Decisions in Academia, *Educational Management Administration & Leadership*, 38, 472-486.

RODOLFA, E.R.; VIEILLE, R.; RUSSELL, P.; NIJJER, S.; NGUYEN, D.Q.; MENDOZA, M.; PERRIN, L. (1999). Internship Selection – Inclusion and Exclusion Criteria, *Professional Psychology: Research and Practice,* 30, 415-419.

ROOTH, D.-O. (2009). Obesity, Attractiveness, and Differential Treatment in Hiring – A Field Experiment, *Journal of Human Resources*, 44, 710-735.

ROSE, G.L.; ANDIAPPAN, P. (1978). Sex Effects on Managerial Hiring Decisions, *Academy of Management Journal*, 21, 104-112.

ROSEN, B.; JERDEE, T.H. (1976a). The Nature of Job-Related Age Stereotypes, *Journal of Applied Psychology*, 61, 180-183.

ROSEN, B.; JERDEE, T.H. (1976b). The Influence of Age Stereotypes on Managerial Decisions, *Journal of Applied Psychology*, 61, 428-432.

ROSENFELD, P. (1997). Impression Management, Fairness, and the Employment Interview, *Journal of Business Ethics*, 16, 801-808.

ROSS, S.A. (1973). The Economic Theory of Agency – The Principal's Problem, *American Economic Review: Papers and Proceedings*, 63, 134-139.

ROTH, P.L.; BEVIER, C.A.; SWITZER III, F.S.; SCHIPPMANN, J.S. (1996). Meta-Analyzing the Relationship between Grades and Job Performance, *Journal of Applied Psychology*, 81, 548-556.

ROTH, P.L.; BOBKO, P. (2000). College Grade Point Average as a Personnel Selection Device – Ethnic Group Differences and Potential Adverse Impact, *Journal of Applied Psychology*, 85, 399-406.

RUFFLE, B.J.; SHTUDINER, Z. (2011). Are Good-Looking People More Employable, Working Paper October 2011.

RYNES, S.; GERHART, B. (1990). Interviewer Assessments of Applicant 'Fit' – An Exploratory Investigation, *Personnel Psychology*, 43, 13-35.

RYNES, S.L.; ORLITZKY, M.O.; BRETZ, R.D. JR. (1997). Experienced Hiring versus College Recruiting – Practices and Emerging Trends. *Personnel Psychology*, 50, 309-339.

SACCO, J.M.; SCHEU, C.R.; RYAN, A.M.; SCHMITT, N. (2003). An Investigation of Race and Sex Similarity Effects in Interviews – A Multilevel Approach to Relational Demography, *Journal of Applied Psychology*, 88, 852-865.

SADOWSKI, D. (2002). Personalökonomie und Arbeitspolitik, Schäffer-Poeschel, Stuttgart.

SALGADO, J.F. (1997). The Five Factor Model of Personality and Job Performance in the European Community, *Journal of Applied Psychology*, 82, 30-43.

SARCLETTI, A. (2010). The Effect of Practical Experience during Studies on Labour Market Entry of Bavarian Graduates, Presentation at the Conference *"Higher Education and beyond – Inequalities Regarding Entrance to Higher Education and Educational Credentials"* Ascona, 8th July 2010.

SARCLETTI, A. (2009). Die Bedeutung von Praktika und studentischen Erwerbstätigkeiten für den Berufseinstieg, *Studien zur Hochschulforschung* 77, München.

SCHICK, G.J.; KUNNECKE, B.F. (1981). Do High Grades, Top Schools, or an Advanced Degree Lead to Job Security and Extraordinary Salary Progression? *Interfaces,* 11, 9-18.

SCHMID MAST, M.; BANGERTER, A.; BULLIARD, C.; AERNI, G. (2011). How Accurate Are Recruiters' First Impressions of Applicants in Employment Interviews? *International Journal of Selection and Assessment*, 19, 198-208.

SCHMIDT, N. (2009). Bundesausbildungsförderung von den Anfängen bis 2007, *Statistisches Bundesamt – Wirtschaft und Statistik 2/2009*, 157-168.

SCHMIDT, F.L.; HUNTER, J.E.; OUTERBRIDGE, A.N. (1986). Impact of Job Experience and Ability on Job Knowledge, Work Sample Performance, and Supervisory Ratings of Job Performance, *Journal of Applied Psychology,* 71, 432-439.

SCHNAPP, K.-U. (1997). Soziodemographische Merkmale der bundesdeutschen Eliten, in Bürklin, W.; Rebenstorf, H. et al. (Editors): Eliten in Deutschland – Rekrutierung und Integration, Leske + Budrich, Opladen, 101-121.

SCHULER, H.; FRIER, D.; KAUFFMANN, M. (1993). Personalauswahl im europäischen Vergleich, Beiträge zur Organisationspsychologie Band 13, Verlag für Angewandte Psychologie, Göttingen.

SCHULER, H.; HÖFT, S. (2004). Diagnose beruflicher Eignung und Leistung, in Schuler, H. (Editor): Lehrbuch Organisationspsychologie, 3rd Ed., Verlag Hans Huber, Bern, 289-344.

SHRM/SOCIETY FOR HUMAN RESOURCE MANAGEMENT (2005). 2004 Reference and Background Checking Survey Report, SHRM, Alexandria, VA.

SILLES, M.A. (2010). Personality, Education and Earnings, *Education Economics*, 18, 131-151.

SINGELL, L.D. JR. (2002). Merit, Need, and Student Self Selection – Is there a Discretion in the Packaging of Aid at a Large Public University? *Economics of Education Review,* 21, 445-454.

SINGER, M.S.; BRUHNS, C. (1991). Relative Effect of Applicant Work Experience and Academic Qualification on Selection Interview Decisions – A Study of Between-Sample Generalizability, *Journal of Applied Psychology*, 76, 550-559.

SMEETERS, B; NAYER, A. (1998). La Discrimination à l'Accès à l'Emploi en Raison de l'Origine Étrangère – Le Cas de la Belgique. *International Migration Papers*, No. 23, International Labour Office, Geneva.

SPENCE, M. (1973). Job Market Signaling, *Quarterly Journal of Economics,* 87, 355-374.

STANLEY, R.E.; FRENCH, P.E. (2009). Evaluating Increased Enrollment Levels in Institutions of Higher Education – A Look at Merit-Based Scholarship Programs, *Public Administration Quarterly*, 33, 4-36.

STEVENS, C.K.; KRISTOF, A.L. (1995). Making the Right Impression – A Field Study of Applicant Impression Management During Job Interviews, *Journal of Applied Psychology*, 80, 587-606.

STONE, C.I.; SAWATZKI, B. (1980). Hiring Bias and the Disabled Interviewee – Effects of Manipulating Work History and Disability Information of the Disabled Job Applicant, *Journal of Vocational Behavior,* 16, 96-104.

STORCK, J.; SPROULL, L. (1995). Through a Glass Darkly – What Do People Learn in Videoconferences? *Human Communication Research*, 22, 197-219.

STRINGER, N. (2008). Aptitude Tests versus School Exams as Selection Tools for Higher Education and the Case for Assessing Educational Achievement in Context, *Research Papers in Education,* 23, 53-68.

SWIDER, B.W.; BARRICK, M.R.; HARRIS, T.B.; STOVERINK, A.C. (2011). Managing and Creating an Image in the Interview – The Role of Interviewee Initial Impressions, *Journal of Applied Psychology*, 96, 1275-1288.

TAJFEL, H. (1982). Social Identity and Intergroup Relations, Cambridge University Press, New York, NY.

TAJFEL, H.; TURNER, J. (1979). An Integrative Theory of Intergroup Conflict, in Austin, W.G.; Worchel, S. (Editors): The Social Psychology of Intergroup Relations, Brooks/Cole, Monterey, CA, 33-48.

TAY, C.; ANG, S.; VAN DYNE, L. (2006). Personality, Biographical Characteristics, and Job Interview Success – A Longitudinal Study of the Mediating Effects of Interviewing Self-Efficacy and the Moderating Effects of Internal Locus of Causality, *Journal of Applied Psychology*, 91, 446-454.

TAYLOR, M.S. (1988). Effects of College Internships on Individual Participants, *Journal of Applied Psychology,* 73, 393-401.

TEDESCHI, J. T.; MELBURG, V. (1984). Impression Management and Influence in the Organization, in Bacharach, S. B.; Lawler, E. J. (Editors): Research in the Sociology of Organizations, Vol. 3, JAI, Greenwich, CT, 31-58.

TETT, R.P.; JACKSON, D.N.; ROTHSTEIN, M. (1991). Personality Measures as Predictors of Job Performance – A Meta-Analytic Review, *Personnel Psychology*, 44, 703-742.

TEWS, M.J.; STAFFORD, K.; ZHU, J. (2009). Beauty Revisited – The Impact of Attractiveness, Ability, and Personality in the Assessment of Employment Suitability, *International Journal of Selection and Assessment,* 17*,* 92-100.

THOMS, P.; MCMASTERS, R.; ROBERTS, M.R.; DOMBKOWSKI, D.A. (1999). Résumé Characteristics as Predictors of an Invitation to Interview, *Journal of Business and Psychology*, 13, 339-356.

TOMMASI, G.W.; WILLIAMS, K.B.; NORDSTROM, C.R. (1998). Letters of Recommendation – What Information Captures HR Professionals' Attention? *Journal of Business and Psychology,* 13, 5-18.

TSAI, W.-C.; CHEN, C.-C.; CHIU, S.-F. (2005). Exploring Boundaries of the Effects of Applicant Impression Management Tactics in Job Interviews, *Journal of Management*, 31, 108-125.

TSAI, W.-C.; CHI, N.-W.; HUANG, T.-C.; HSU, A.-J. (2011). The Effects of Applicant Résumé Contents on Recruiters' Hiring Recommendations – The Mediating Roles of Recruiter Fit Perceptions, *Applied Psychology: An International Review,* 60, 231-254.

TURBAN, D.B.; JONES, A.P. (1988). Supervisor-Subordinate Similarity – Types, Effects, and Mechanisms, *Journal of Applied Psychology*, 73, 228-234.

TURNER, J.C. (1987). Rediscovering the Social Group – A Self-Categorization Theory, Basil Blackwell, Oxford.

TURNER, M.A.; FIX, M.; STRUYK, R.J. (1991). Opportunities Denied, Opportunities Diminished – Racial Discrimination in Hiring, Urban Institute Report 91-9, The Urban Institute Press, Washington DC.

TURNER, R.; NICHOLSON, S. (2011). Reasons Selectors Give for Accepting and Rejecting Medical Applicants before Interview, *Medical Education,* 45, 298-307.

UGBAH, S.D.; MAJORS, R.E. (1992). Influential Communication Factors in Employment Interviews, *The Journal of Business Communication*, 29, 145-159.

URBAN, D. (1993). Logit-Analyse – Statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen, Gustav Fischer, Stuttgart.

URLINGS-STROP,L.C.; STIJNEN, T.; THEMMEN, A.P.N.; SPLINTER, T.A.W. (2009). Selection of Medical Students – A Controlled Experiment, *Medical Education,* 43, 175-183.

VAN AUER, L. (2007). Ökonometrie – Eine Einführung, 4th Ed., Springer-Verlag, Berlin.

VAN IDDEKINGE, C.H.; MCFARLAND, L.A.; RAYMARK, P.H. (2007). Antecedents of Impression Management Use and Effectiveness in a Structured Interview? *Journal of Management*, 33, 752-773.

VAN OURS, J.; RIDDER, G. (1991). Job Requirements and the Recruitment of New Employees, *Economics Letters,* 36, 213-218.

VARMA, A.; TOH, S.M.; PICHLER, S. (2006). Ingratiation in Job Applications – Impact on Selection Decisions, *Journal of Managerial Psychology*, 21, 200-210.

VEALL, M.R.; ZIMMERMANN, K.F. (1996). Pseudo-R^2 Measures for Some Common Limited Dependent Variable Models, *Journal of Economic Surveys*, 10, 241-259.

VECCHIO, R.P.; BULLIS, R.C. (2001). Moderators of the Influence of Supervisor-Subordinate Similarity on Subordinate Outcomes, *Journal of Applied Psychology*, 86, 884-896.

VGRDL/Volkswirtschaftliche Gesamtrechnungen der Länder (2011) – Arbeitnehmerentgelt, Bruttolöhne und -gehälter in den Ländern und Ost-West-Großraumregionen Deutschlands 1991 bis 2010, Reihe 1, Länderergebnisse Band 2, Berechnungsstand August 2010/Februar 2011, Statistische Ämter der Länder, Frankfurt a. M.

VON ROSENBLADT, B.; THUM, M.; WASCHBÜSCH, E.; KIENDL, B. (1995). Gesamtdeutsche Elitestudie 1995 – Methodenbericht zur Durchführung der Befragung, Infratest Burke Sozialforschung, München.

WALLACH, M.A.; KOGAN, N. (1959). Sex Differences and Judgment Processes, *Journal of Personality*, 27, 555-564.

WALSH, J.P.; WEINBERG, R.M.; FAIRFIELD, M.L. (1987). The Effects of Gender on Assessment Centre Evaluations, *Journal of Occupational Psychology*, 60, 305-309.

WANNER, R.A.; MCDONALD, L. (1983). Ageism in the Labor Market – Estimating Earnings Discrimination against Older Workers, *Journal of Gerontology*, 38, 738-744.

WARD, M. (2001). The Gender Salary Gap in British Academia, *Applied Economics,* 33, 1669-1681.

WARR, P. (1994). Age and Job Performance, in Snel, J.; Cremer, R. (Editors): Work and Aging – A European Perspective, Taylor & Francis, London, 309-322.

WARR, P.B.; KNAPPER, C. (1968). The Perception of People and Events, Wiley, London.

WATKINS, L.M.; JOHNSTON, L. (2000). Screening Job Applicants – The Impact of Physical Attractiveness and Application Quality, *International Journal of Selection and Assessment,* 8, 76-84.

WEICHSELBAUMER, D. (2004). Is it Sex or Personality? The Impact of Sex Stereotypes on Discrimination in Applicant Selection, *Eastern Economic Journal,* 30, 159-186.

WESTPHAL, J.D.; ZAJAC, E.J. (1995). Who Shall Govern? CEO/Board Power, Demographic Similarity, and New Director Selection, *Administrative Science Quarterly*, 40, 60-83.

WEXLEY, K.N.; SANDERS, R.E.; YUKL, G.A. (1973). Training Interviewers to Eliminate Contrast Effects in Employment Interviews, *Journal of Applied Psychology*, 57, 233-236.

WEXLEY, K.N.; YUKL, G.A.; KOVACS, S.Z.; SANDERS, R.E. (1972). Importance of Contrast Effects in Employment Interviews, *Journal of Applied Psychology*, 56, 45-48.

WILDENMANN, R.; KAASE, M.; HOFFMANN-LANGE, U.; KUTTEROFF, A.; WOLF, G. (1982). Führungsschicht in der Bundesrepublik Deutschland 1981, Universität Mannheim.

WINGROVE, J.; GLENDINNING, R.; HERRIOT, P. (1984). Graduate Pre-Selection – A Research Note, *Journal of Occupational Psychology,* 57, 196-171.

WISSENSCHAFTSRAT (2007). Prüfungsnoten im Prüfungsjahr 2005 – an Universitäten (einschließlich KH, PH, TH) sowie an Fachhochschulen (einschließlich Verwaltungshochschulen) nach ausgewählten Studienbereichen und Studienfächern, Arbeitsbericht, Köln.

WOOLDRIDGE, J.M. (2009). Introductory Econometrics – A Modern Approach, 4th Ed., International Student Edition, South-Western Cengage Learning.

ZIKMUND, W.G.; HITT, M.A.; PICKENS, B.A. (1978). Influence of Sex and Scholastic Performance on Reactions to Job Applicant Résumés, *Journal of Applied Psychology,* 63, 252-254.

ZIMMERMANN, K.F. (2010). Deutschland braucht auf Dauer Fachkräfte – auch durch Zuwanderung, *Wochenbericht des Deutschen Instituts für Wirtschaftsforschung (DIW)*, 77 (46), 28.

ZINOVYEVA, N.; BAGUES, M. (2011). Does Gender Matter for Academic Promotion? Evidence from a Randomized Natural Experiment, *IZA Working Paper* No. 5537.

LEGISLATIVE TEXTS

AGE DISCRIMINATION IN EMPLOYMENT ACT (1967) as amended, as it appears in Vol. 29 of the United States Code, provided by the US Equal Employment Opportunity Commission (EEOC), accessed October 24, 2012 at: http://www.eeoc.gov/laws/statutes/adea.cfm

AGG/ALLGEMEINES GLEICHBEHANDLUNGSGESETZ (2006) as provided by the Bundesministerium der Justiz, accessed October 24, 2012 at: http://www.gesetze-im-internet.de/agg/BJNR189710006.html

AMERICANS WITH DISABILITIES ACT (1990) as amended in 2008, provided by the US Department of Justice, accessed October 24, 2012 at: http://www.ada.gov/pubs/ada.htm

EU DIRECTIVE 2002/73/EC (2002) amending Council Directive 76/207/EEC on the Implementation of the Principle of Equal Treatment for Men and Women as Regards Access to Employment, Vocational Training and Promotion, and Working Conditions, *Official Journal of the European Communities*, L 269, 15-20, accessed October 24, 2012 at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0073:EN:PDF

EU DIRECTIVE 2000/43/EC (2000) implementing the Principle of Equal Treatment between Persons Irrespective of Racial or Ethnic Origin, *Official Journal of the European Communities*, L 180, 22-26, accessed October 24, 2012 at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0043:en:HTML

EU DIRECTIVE 2000/78/EC (2000) establishing a General Framework for Equal Treatment in Employment and Occupation, *Official Journal of the European Communities*, L 303, 16-22, accessed October 24, 2012 at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0078:en:HTML

FEDERAL CIVIL RIGHTS ACT (1964), Online Version accessed October 24, 2012 at: http://www.ourdocuments.gov/doc.php?flash=true&doc=97&page=transcript

APPENDIX

APPENDIX 1: DESCRIPTIVE STATISTICS – 'NO SHOWS' VS. 'INTERVIEWED APPLICANTS'

Variable	Obs.	Mean	SD	Min	Max	Sample
DEPENDENT VARIABLES						
Pre-Selection Score	254	84.83	5.37	73	100	Invited & Interviewed
	18	86.53	4.78	78	94	Invited, but did not show up
INDEPENDENT VARIABLES						
Academic Achievement						
High School Grade Average	254	1.56	.466	1	3.3	Invited & Interviewed
	18	1.52	.392	1.1	2.3	Invited, but did not show up
(Preliminary) University Grade Average	248	1.69	.445	1	3.5	Invited & Interviewed
	18	1.78	.451	1.07	2.64	Invited, but did not show up
Field of Studies						
Engineering	254	.142	-	0	1	Invited & Interviewed
	18	.167	-	0	1	Invited, but did not show up
Mathematics, Informatics and Natural Sciences	254	.185	-	0	1	Invited & Interviewed
	18	.167	-	0	1	Invited, but did not show up
Law, Economics and Social Sciences	254	.350	-	0	1	Invited & Interviewed
	18	.333	-	0	1	Invited, but did not show up
Linguistic and Cultural Sciences	254	.295	-	0	1	Invited & Interviewed
	18	.278	-	0	1	Invited, but did not show up
Other	254	.028	-	0	1	Invited & Interviewed
	18	.056	-	0	1	Invited, but did not show up
Type of Home Institution						
University of Applied Sciences	253	.079	-	0	1	Invited & Interviewed
	18	.056	-	0	1	Invited, but did not show up
Technical University	253	.154	-	0	1	Invited & Interviewed
	18	.278	-	0	1	Invited, but did not show up
University	253	.767	-	0	1	Invited & Interviewed
	18	.667	-	0	1	Invited, but did not show up
Duration of Study						
Semester	254	4.41	1.48	2	14	Invited & Interviewed
	18	4.22	1.77	2	14	Invited, but did not show up
Work Experience Part-Time						
(Previous or Current) Part-Time Job only at University	254	.181	-	0	1	Invited & Interviewed
	18	.222	-	0	1	Invited, but did not show up
(Previous or Current) Part-Time Jobs both at & outside University	254	.252	-	0	1	Invited & Interviewed
	18	.222	-	0	1	Invited, but did not show up
(Previous or Current) Part-Time Job only outside University	254	.358	-	0	1	Invited & Interviewed
	18	.389	-	0	1	Invited, but did not show up
No Part-Time Job	254	.209	-	0	1	Invited & Interviewed
	18	.167	-	0	1	Invited, but did not show up
Other Scholarships						
'BAföG'	254	.181	-	0	1	Invited & Interviewed
	18	.111	-	0	1	Invited, but did not show up
Previous Merit-Based Stipend	254	.232	-	0	1	Invited & Interviewed
	18	.333	-	0	1	Invited, but did not show up
'Studienstiftung des Dt. Volkes'	254	.102	-	0	1	Invited & Interviewed
	18	.167	-	0	1	Invited, but did not show up
Also Applied for Other Scholarships	254	.492	-	0	1	Invited & Interviewed
	18	.556	-	0	1	Invited, but did not show up

Number of Extracurricular Activities (ECAs)						
One Type of ECAs	254	.260	-	0	1	Invited & Interviewed
	18	.389	-	0	1	Invited, but did not show up
More than one Type of ECAs	254	.587	-	0	1	Invited & Interviewed
	18	.611	-	0	1	Invited, but did not show up
No Extracurricular Activity	254	.154	-	0	1	Invited & Interviewed
	18	.000	-	0	1	Invited, but did not show up
Letter of Reference						
Recommendation	246	9.46	.493	8	10	Invited & Interviewed
	17	9.40	.545	7.75	10	Invited, but did not show up
Recommending Person=Professor	254	.744	-	0	1	Invited & Interviewed
	18	.611	-	0	1	Invited, but did not show up
Good Relation to Recommending Person	241	.905	-	0	1	Invited & Interviewed
	18	.889	-	0	1	Invited, but did not show up
Language Proficiency						
Language Skills	253	90.58	7.35	59.2	100	Invited & Interviewed
	17	90.07	8.43	70	100	Invited, but did not show up
TOEFL	254	.756	-	0	1	Invited & Interviewed
	18	.389	-	0	1	Invited, but did not show up
Project-Specific Statements						
Tuition Fees (in Th. $US)	248	15.626	14.28	0	80	Invited & Interviewed
	18	20.471	18.52	0	51.81	Invited, but did not show up
Private Guest Institution	254	.283	-	0	1	Invited & Interviewed
	18	.222	-	0	1	Invited, but did not show up
Top10 Guest Institution	254	.181	-	0	1	Invited & Interviewed
	18	.278	-	0	1	Invited, but did not show up
Guest Institution in Canada	254	.213	-	0	1	Invited & Interviewed
	18	.167	-	0	1	Invited, but did not show up
Guest Institution in US-Midwest	254	.118	-	0	1	Invited & Interviewed
	18	.167	-	0	1	Invited, but did not show up
Guest Institution in US-Northeast	254	.252	-	0	1	Invited & Interviewed
	18	.111	-	0	1	Invited, but did not show up
Guest Institution in US-South	254	.181	-	0	1	Invited & Interviewed
	18	.222	-	0	1	Invited, but did not show up
Guest Institution in US-West	254	.236	-	0	1	Invited & Interviewed
	18	.333	-	0	1	Invited, but did not show up
Applicant Characteristics						
Gender	254	.492	-	0	1	Invited & Interviewed
	18	.556	-	0	1	Invited, but did not show up
Glasses	253	.213	-	0	1	Invited & Interviewed
	18	.222	-	0	1	Invited, but did not show up
Born in Germany	254	.945	-	0	1	Invited & Interviewed
	18	.945	-	0	1	Invited, but did not show up
Age	254	21.67	1.18	19	27	Invited & Interviewed
	18	21.22	1.44	19	24	Invited, but did not show up
Parents=Academics	254	.201	-	0	1	Invited & Interviewed
	18	.167	-	0	1	Invited, but did not show up
Pre-Selection Evaluator Characteristics and Pre-Selection Situation						
Evaluator Gender	254	.205	-	0	1	Invited & Interviewed
	18	.278	-	0	1	Invited, but did not show up
Evaluator Age	253	53.74	9.76	33	68	Invited & Interviewed
	18	51.06	8.07	40	68	Invited, but did not show up
Year						
Selection Year	254	.508	-	0	1	Invited & Interviewed
	18	.667	-	0	1	Invited, but did not show up

APPENDIX 2: PROBIT REGRESSION RESULTS

Probit Regression Coefficients		**Estimation Number**				
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**P - I**	**P - IIa**	**P - IIb**	**P - IIc**	**P - III**
Academic Achievement	High School Grade Average	**-0.708*** **	**-0.614*** **	**-0. 577*** **	**-0.560*** **	**-0.545*** **
	University Grade Average	**-1. 012*** **	**-0.990*** **	**-0.954*** **	**-0.972*** **	**-0.922*** **
Home Institution Characteristics	*Reference Category: University*					
	University of Applied Sciences	-/-	-0.571	-0.516	-0.386	-0.385
	Technical University	-/-	-0.299	-0.335	-0.316	-0.285
Field of Study	*Reference Category: Law, Economics and Social Sciences*					
	Engineering	-/-	-/-	-/-	0.325	-/-
	Mathematics, Informatics & Natural Sciences	-/-	-/-	-/-	0.158	-/-
	Linguistic and Cultural Sciences	-/-	-/-	-/-	0.139	-/-
	Other Fields of Study	-/-	-/-	-/-	0.484	-/-
Language Proficiency	Language Skills	-/-	**0.019***	**0.019***	**0.020***	0.018
	TOEFL	-/-	**0.487*** **	**0.542*** **	**0.562*** **	**0.513*** **
Work Experience Part-Time	*Reference Category: No Part-Time Job*					
	at University	-/-	0.126	0.123	0.150	0.159
	both at University & outside University	-/-	0.039	0.074	0.106	0.133
	outside University	-/-	-0.066	-0.033	-0.002	0.010
Recipient of other Sponsorships	'BAföG'	-/-	0.055	0.054	0.082	0.039
	Any other Merit-Based Stipend	-/-	0.147	-/-	-/-	-/-
	'Studienstiftung des Deutschen Volkes'	-/-	-/-	**0.662** **	**0.662** **	**0.663** **
	Applied for other Scholarships	-/-	0.198	0.199	0.178	0.238
Project-Specific Statements	Tuition Fees (in Thousand US)	-/-	-0.002	-0.003	-0.002	-0.003
	Private Guest Institution	-/-	0.161	0.170	0.210	0.203
	Guest Institution=Top10 University	-/-	0.028	-0.021	-0.036	0.006
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*					
	Canada	-/-	0.052	0.076	0.063	0.138
	US-Midwest	-/-	0.285	0.266	0.242	0.310
	US-North-East	-/-	-0.169	-0.186	-0.212	-0.163
	US-South	-/-	0.082	0.094	0.062	0.010
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*					
	One Type of ECAs	-/-	**0.593** **	**0.569** **	**0.547** **	**0.569** **
	More Types of ECAs	-/-	**0.585** **	**0.555** **	**0.562** **	**0.566** **
Letter of Reference	Recommendation	-/-	0.119	0.123	0.129	0.122
	Recommending Person=Professor	-/-	**0.566*** **	**0.553*** **	**0.560*** **	**0.587*** **
	Good Relation to Recommending Person	-/-	**0.414***	**0.410***	0.400	0.406
Applicant Characteristics	Gender	-/-	-/-	-/-	-/-	0.042
	Glasses	-/-	-/-	-/-	-/-	0.173
	Born in Germany	-/-	-/-	-/-	-/-	0.501
	Age	-/-	-/-	-/-	-/-	**-0.159** **
	Parents=Academics	-/-	-/-	-/-	-/-	-0.104
Control Variables	*Semester*	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Participant in Organized Study-Abroad Program	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Selection Year	-/-	*incl.*	*incl.*	*incl.*	*incl.*
Constant		**-0.634*** **	**-2.442*** **	**-2.461*** **	**-2.613*** **	**-3.047*** **
Observations		429	429	429	429	429
Pseudo R^2		0.191	0.289	0.296	0.300	0.310
*significant at * p <0.10 ** p <0.05 *** p <0.01*						

Probit Regression Coefficients		Estimation Number					
Dependent Variable: Invitation (Yes=1 No=0)		**P - IV**	**P - Va**	**P - Vb**	**P - VI**	**P - VIIa**	**P - VIIb**
Academic Achievement	High School Grade Average	**-0.838*** **	**-0.640*** **	**-0.671*** **	**-0.673*** **	**-0.767*** **	**-0.880*** **
	University Grade Average	**-0.974*** **	**-0.831*** **	**-0.792*** **	**-0.828*** **	**-0.805*** **	**-1.863*** **
Home Institution Characteristics	*Reference Category: University*						
	University of Applied Sciences	-/-	**-0.814** **	**-0.796** **	-0.641	-0.377	-0.140
	Technical University	-/-	-0.365	**-0.414**	-0.306	-0.171	-0.173
Field of Study	*Reference Category: Law, Economics and Social Sciences*						
	Engineering	-/-	-/-	-/-	-/-	-/-	0.540
	Mathematics, Informatics & Natural Sciences	-/-	-/-	-/-	-/-	-/-	-0.158
	Linguistic and Cultural Sciences	-/-	-/-	-/-	-/-	-/-	**0.494***
	Other Fields of Study	-/-	-/-	-/-	-/-	-/-	0.295
Language Proficiency	Language Skills	-/-	**0.028*** **	**0.028*** **	**0.026** **	**0.024** **	**0.023** **
	TOEFL	-/-	**0.808*** **	**0.687*** **	**0.821*** **	**0.806*** **	**0.839*** **
Work Experience Part-Time	*Reference Category: No Part-Time Job*						
	at University	-/-	-0.062	-0.056	0.004	-0.043	-0.055
	both at University & outside University	-/-	0.414	**0.455***	**0.549***	**0.619** **	**0.626** **
	outside University	-/-	-0.189	-0.153	-0.149	-0.230	-0.205
Recipient of other Sponsorships	'BAföG'	-/-	-0.231	-0.220	-0.268	-0.353	-0.335
	Any other Merit-Based Stipend	-/-	**0.533** **	-/-	-/-	-/-	-/-
	'Studienstiftung des Dt. Volkes'	-/-	-/-	**1.301** **	**1.227** **	**1.115***	**1.200** **
	Applied for other Scholarships	-/-	-0.097	-0.096	-0.050	-0.028	-0.016
Project-Specific Statements	Tuition Fees (in Thousand US)	-/-	0.009	0.010	0.009	0.007	0.010
	Private Guest Institution	-/-	**-0.428***	**-0.413***	-0.423	**-0.520***	**-0.502***
	Guest Institution=Top10 University	-/-	**-0.401***	**-0.423***	-0.379	-0.390	-0.335
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*						
	Canada	-/-	0.003	0.041	0.020	0.081	0.159
	US-Midwest	-/-	-0.046	-0.062	0.029	0.077	0.037
	US-Northeast	-/-	0.409	0.443	**0.506***	**0.611** **	**0.619** **
	US-South	-/-	0.183	0.214	0.235	0.247	0.218
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*						
	One Type of ECAs	-/-	**0.519** **	**0.549** **	**0.516***	**0.605** **	**0.616** **
	More Types of ECAs	-/-	**0.603** **	**0.607*** **	**0.616** **	**0.711*** **	**0.710*** **
Letter of Reference	Recommendation	-/-	**0.339** **	**0.348** **	**0.342** **	**0.387** **	**0.395*** **
	Recommending Person=Professor	-/-	**0.570*** **	**0.561*** **	**0.659*** **	**0.726*** **	**0.814*** **
	Good Relation to Recommender	-/-	**0.828*** **	**0.872*** **	**0.926*** **	**0.941*** **	**0.854*** **
Applicant Characteristics	Gender	-/-	-/-	-/-	0.124	0.069	0.044
	Glasses	-/-	-/-	-/-	0.231	0.267	0.275
	Born in Germany	-/-	-/-	-/-	**0.976*** **	**1.010*** **	**0.977*** **
	Age	-/-	-/-	-/-	**-0.213** **	**-0.236** **	**-0.241*** **
	Parents=Academics	-/-	-/-	-/-	**-0.400** **	**-0.459** **	**-0.492** **
Rater Characteristics	Evaluator Gender	-/-	-/-	-/-	-/-	**0.785*** **	**0.851*** **
	Evaluator Age	-/-	-/-	-/-	-/-	**0.021***	**0.039***
Control Variables	*Semester*	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
	Part. in Organized Study-Abroad Program	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
	Selection Year	-/-	*incl.*	*incl.*	*incl.*	*incl.*	*incl.*
Constant		**0.214*** **	**-1.872*** **	**-1.957*** **	**-2.971*** **	**-3.328*** **	**-3.520*** **
Observations		423	423	423	423	423	423
Pseudo R^2		0.226	0.418	0.422	0.457	0.473	0.488

*significant at * p <0.10 ** p <0.05 *** p <0.01*

Probit Regression Coefficients		Estimation Number				
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**P - VIII**	**P - IXa**	**P - IXb**	**P - Xa**	**P - Xb**
Academic Achievement	High School Grade Average	**-0.451****	-0.353	-0.290	-0.216	-0.183
	University Grade Average	**-0.656*****	**-0.929*****	**-0.900*****	**-0.890*****	**-0.973*****
Home Institution Characteristics	*Reference Category: University*					
	University of Applied Sciences	-/-	-0.520	-0.413	-0.287	-0.231
	Technical University	-/-	-0.020	-0.015	0.116	0.138
Field of Study	*Reference Category: Law, Economics and Social Sciences*					
	Engineering	-/-	-/-	-/-	-/-	0.234
	Mathematics, Informatics & Natural Sciences	-/-	-/-	-/-	-/-	0.255
	Linguistic and Cultural Sciences	-/-	-/-	-/-	-/-	-0.149
	Other Fields of Study	-/-	-/-	-/-	-/-	**1.561***
Language Proficiency	Language Skills	-/-	0.013	0.012	0.012	0.017
	TOEFL	-/-	0.007	0.113	0.079	0.056
Work Experience Part-Time	*Reference Category: No Part-Time Job*					
	at University	-/-	0.326	0.340	0.343	0.493
	both at University & outside University	-/-	-0.050	0.002	0.049	0.116
	Outside University	-/-	0.081	0.150	0.138	0.209
Recipient of other Sponsorships	'BAföG'	-/-	0.087	0.080	0.028	0.095
	Any other Merit-Based Stipend	-/-	0.004	-/-	-/-	-/-
	'Studienstiftung des Deutschen Volkes'	-/-	-/-	0.733	**0.826***	**0.845***
	Applied for other Scholarships	-/-	**0.430***	**0.415***	**0.511****	**0.538****
Project-Specific Statements	Tuition Fees (in Thousand US)	-/-	-0.010	-0.010	-0.010	-0.012
	Private Guest Institution	-/-	0.406	0.437	0.529	0.525
	Guest Institution=Top10 University	-/-	0.353	0.272	0.330	0.301
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*					
	Canada	-/-	-0.066	-0.036	0.067	0.011
	US-Midwest	-/-	0.557	0.507	0.598	0.621
	US-Northeast	-/-	-0.532	-0.604	-0.560	-0.613
	US-South	-/-	-0.114	-0.128	-0.122	-0.192
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*					
	One Type of ECAs	-/-	**0.670****	**0.636***	**0.717****	**0.744****
	More Types of ECAs	-/-	**0.569***	**0.515***	**0.574***	**0.665****
Letter of Reference	Recommendation	-/-	-0.176	-0.178	-0.210	-0.223
	Recommending Person=Professor	-/-	0.300	0.287	0.325	0.264
	Good Relation to Recommending Person	-/-	-0.319	-0.315	-0.301	-0.315
Applicant Characteristics	Gender	-/-	-/-	-/-	0.226	0.250
	Glasses	-/-	-/-	-/-	0.251	0.165
	Born in Germany	-/-	-/-	-/-	0.160	0.215
	Age	-/-	-/-	-/-	**-0.178***	-0.168
	Parents=Academics	-/-	-/-	-/-	0.078	0.157
Control Variables	*Semester*	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Participant in Organized Study-Abroad Program	-/-	*incl.*	*incl.*	*incl.*	*incl.*
	Selection Year	-/-	*incl.*	*incl.*	*incl.*	*incl.*
Constant		**0.250*****	-0.585	-0.677	-1.269	**-1.375***
Observations		226	226	226	226	226
Pseudo R^2		0.069	0.169	0.179	0.196	0.217
*significant at * p <0.10 ** p <0.05 *** p <0.01*						

Probit Regression Coefficients		Estimation Number		
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**P - XI**	**P - XII**	**P - XIII**
Academic Achievement	High School Grade Average	-0.319	-0.297	-0.083
	University Grade Average	**-0.920*****	**-1.065*****	**-1.046*****
Language Proficiency	Language Skills	0.010	0.013	0.012
Recipient of other Sponsorships	'Studienstiftung des Deutschen Volkes'	**0.740***	**1.024****	**1.011****
	Applied for other Scholarships	**0.419****	**0.522****	**0.481****
Project-Specific Statements	Tuition Fees (in Thousand US)	-0.011	-0.012	-0.012
	Private Guest Institution	0.079	0.124	0.129
	Guest Institution=Top10 University	0.257	0.257	0.193
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*			
	One Type of ECAs	**0.588***	**0.572***	**0.626***
	More Types of ECAs	0.483	0.519	0.534
Letter of Reference	Recommendation	-0.183	-0.252	-0.348
	Recommending Person=Professor	0.302	0.193	0.138
	Good Relation to Recommending Person	-0.225	-0.234	-0.253
Applicant Characteristics	Gender	0.180	0.156	0.154
	Born in Germany	0.204	0.479	0.471
	Age	**-0.172***	-0.166	**-0.182***
	Parents=Academics	0.112	0.197	0.187
Evaluation Committee Characteristics	Size	-0.030	-0.134	-0.139
	Fraction of Female Evaluators	0.953	0.775	0.659
	Mainly Male Evaluators	0.218	0.316	0.209
	Average Evaluator Age	-0.025	-0.026	-0.032
	Dispersion Evaluator Age	-0.007	0.004	0.011
Interview Framework	*Reference Category: Interview Time 10-10:59 a.m.*			
	Interview Time: 09-09:59 a.m.	-/-	-0.013	-0.029
	Interview Time: 11-11:59 a.m.	-/-	**-0.959****	**-0.931****
	Interview Time: 12-01:30 p.m.	-/-	-0.157	-0.195
	Interview Time: 02-02:59 p.m.	-/-	0.251	0.264
	Interview Time: 03-03:59 p.m.	-/-	-0.657	-0.687
	Interview Time: 04-04:59 p.m.	-/-	0.152	0.186
	Interview Time: 05-06:59 p.m.	-/-	-0.227	-0.215
Pre-selection Outcome	Pre-Selection Score	-/-	-/-	**0.074*****
Year	Selection Year	**-0.389***	**-0.479****	**-0.503****
Control Variables	*Semester*	*incl.*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	*incl.*	*incl.*	*incl.*
Constant		-1.006	-1.065	-0.852
Observations		226	226	226
Pseudo R^2		0.176	0.226	0.261
*significant at * p <0.10 ** p <0.05 *** p <0.01*				

Probit Regression Coefficients		Estimation Number	
Dependent Variable: Scholarship Awarded (Yes=1 No=0)		**P - III**	**P - XIV**
Academic Achievement	High School Grade Average	**-0.545*****	-0.282
	University Grade Average	**-0.922*****	**-0.629****
Home Institution Characteristics	*Reference Category: University*		
	University of Applied Sciences	-0.385	-0.184
	Technical University	-0.285	-0.169
Language Proficiency	Language Skills	0.018	0.008
	TOEFL	**0.513*****	0.187
Work Experience Part-Time	*Reference Category: No Part-Time Job*		
	at University	0.159	0.208
	both at University & outside University	0.133	-0.020
	Outside University	0.010	0.030
Recipient of other Sponsorships	'BAföG'	0.039	0.108
	'Studienstiftung des Deutschen Volkes'	**0.663****	0.352
	Applied for other Scholarships	0.238	**0.311***
Project-Specific Statements	Tuition Fees (in Thousand US)	-0.003	-0.008
	Private Guest Institution	0.203	0.363
	Guest Institution=Top10 University	0.006	0.149
Guest Institution Region	*Reference Category: Guest Institution Region=US-West*		
	Canada	0.138	0.099
	US-Midwest	0.310	0.308
	US-Northeast	-0.163	-0.308
	US-South	0.010	0.014
Number of Extracurricular Activities	*Reference Category: No Extracurricular Activity*		
	One Type of ECAs	**0.569****	0.383
	More Types of ECAs	**0.566****	0.339
Letter of Reference	Recommendation	0.122	-0.022
	Recommending Person=Professor	**0.587*****	**0.406***
	Good Relation to Recommending Person	0.406	0.078
Applicant Characteristics	Gender	0.042	0.074
	Glasses	0.173	0.115
	Born in Germany	0.501	0.336
	Age	**-0.159****	-0.098
	Parents=Academics	-0.104	0.065
Pre-selection Success	Individual Invitation Probability	-/-	**1.477****
Control Variables	*Semester*	*incl.*	*incl.*
	Existing Contacts to Guest Institution	*incl.*	*incl.*
	Participant in Organized Study-Abroad Program	*incl.*	*incl.*
	Selection Year	*incl.*	*incl.*
Constant		**-3.047*****	**-3.000*****
Observations		429	423
Pseudo R^2		0.310	0.322

*significant at * p <0.10 ** p <0.05 *** p <0.01*

APPENDIX 3: STANDARD APPLICANT CHARACTERISTICS – OVERALL SELECTION

A standard applicant in Estimations I-III (standard applicant 1) has applied for the investigated scholarship program in **2008**. He has achieved a **high school grade average of 1.76** and **university grades averaging 1.89**. He is enrolled at a **university** and studies **Law, Economics or Social Sciences** in his **4th semester**. His certified **language skills** are reported to be **88.85** (on a scale from 0 to 100) and he **has handed in a TOEFL**. In terms of **part-time employment**, he has already **worked for at least one external employer**, i.e. not at the university. The standard applicant (1) does **neither receive BAföG, nor any other merit based stipend** (and consequently he has not been awarded the stipend of the 'Studienstiftung des Deutschen Volkes'). Additionally, he did not or at least **did not indicate to have applied for other scholarship programs** promising to support the study-abroad experience.

The standard applicant (1) wants to spend his **stay abroad** in the academic year **2009/10** at a **public** higher education institution located in the **US-West** which does **not belong to the THE Top 10 Universities** in this specific field. He expects to pay **tuition fees** totaling **US$ 14,643** for the entire stay abroad. He has **not been in contact with** the respective guest institution and does **not take part in an organized study abroad program**

In terms of **extracurricular activities**, the standard applicant (1) is active in **more than one type of ECA**. He has handed in a **recommendation from a professor** who indicates to **know him very well**. The **recommendation score** the standard applicant (1) has achieved is **9.26**.

The standard applicant (1) is **male**, **does not wear glasses, was born in Germany, is 21.84 years old and did not mention his parents in the CV**. Consequently, evaluators don't know whether or not the standard applicant's parents are academics.

In pre-selection, the standard applicant (1) is **evaluated by a male evaluator who is aged 53.61 years** and has managed to achieve a **pre-selection score of 76.64**.

APPENDIX 4: STANDARD APPLICANT CHARACTERISTICS – PRE-SELECTION

A standard applicant in Estimations IV-VIIb (standard applicant 2) has applied for the investigated scholarship program in **2008**. He has achieved a **high school grade average of 1.76** and **university grades averaging 1.89**. He is enrolled at a **university** and studies **Law, Economics or Social Sciences** in his **4th semester**. His certified **language skills** are reported to be **88.76** (on a scale from 0 to 100) and he **has handed in a TOEFL**. In terms of **part-time employment**, he has already **worked for at least one external employer**, i.e. not at the university. The standard applicant (2) does **neither receive BAföG, nor any other merit based stipend** (and consequently he has not been awarded the stipend of the 'Studienstiftung des Deutschen Volkes'). Additionally, he did not or at least **did not indicate to have applied for other scholarship programs** promising to support the study-abroad experience.

The standard applicant (2) wants to spend his **stay abroad** in the academic year **2009/10** at a **public** higher education institution located in the **US-West** which does **not belong to the THE Top 10 Universities** in this specific field. He expects to pay **tuition fees** totaling **US$ 14,652** for the entire stay abroad. He has **not been in contact with** the respective guest institution and does **not take part in an organized study abroad program**

In terms of **extracurricular activities**, the standard applicant (2) is active in **more than one type of ECA**. He has handed in a **recommendation from a professor** who indicates to **know him very well**. The **recommendation score** the standard applicant (2) has achieved is **9.27**.

The standard applicant (2) is **male**, **does not wear glasses, was born in Germany, is 21.85 years old and did not mention his parents in the CV**. Consequently, evaluators don't know whether or not the standard applicant's parents are academics.

In pre-selection, the standard applicant (2) is **evaluated by a male evaluator who is aged 53.61 years** and has managed to achieve a **pre-selection score of 76.71**.

APPENDIX 5: STANDARD APPLICANT CHARACTERISTICS – FINAL SELECTION

A standard applicant in Estimations VIII-XIII (standard applicant 3) has applied for the investigated scholarship program in **2008**. He has achieved a **high school grade average of 1.55** and **university grades averaging 1.68**. He is enrolled at a **university** and studies **Law, Economics or Social Sciences** in his **4th semester**. His certified **language skills** are reported to be **90.82** (on a scale from 0 to 100) and he **has handed in a TOEFL**. In terms of **part-time employment**, he has already **worked for at least one external employer**, i.e. not at the university. The standard applicant (3) does **neither receive BAföG, nor any other merit based stipend** (and consequently he has not been awarded the stipend of the 'Studienstiftung des Deutschen Volkes'). However, he has also **applied for other scholarship programs** promising to support the study-abroad experience.

The standard applicant (3) wants to spend his **stay abroad** in the academic year **2009/10** at a **public** higher education institution located in the **US-Northeast** which does **not belong to the THE Top 10 Universities** in this specific field. He expects to pay **tuition fees** totaling **US$ 16,081** for the entire stay abroad. He has **already been in contact with** the respective guest institution, but does **not take part in an organized study abroad program.**

In terms of **extracurricular activities**, the standard applicant (3) is active in **more than one type of ECA**. He has handed in a **recommendation from a professor** who indicates to **know him very well**. The **recommendation score** the standard applicant (3) has achieved is **9.47**.

The standard applicant (3) is **male**, **does not wear glasses, was born in Germany, is 21.7 years old and did not mention his parents in the CV**. Consequently, evaluators don't know whether or not the standard applicant's parents are academics.

In pre-selection, the standard applicant (3) is **evaluated by a male evaluator who is aged 53.61 years** and has managed to achieve a **pre-selection score of 84.98**. In final selection, the standard applicant (3) has been **interviewed from 10 to 11** am by an **interview panel** consisting of **4.36 evaluators**. On average, these **panel members are 52.41 years old** (**standard deviation 15.96 years**). The panel **mainly** consists of **male members**: The **fraction of female evaluators** in the committee is **0.26**.

APPENDIX 6: DESCRIPTIVE STATISTICS – REDUCED SAMPLES

Variable	Obs.	Mean	SD	Min	Max	Sample
DEPENDENT VARIABLES						
Pre-Selection Score	504	75.42	14.08	19.5	100	Full
	429	76.64	13.19	19.5	100	Subjects included in Estimations I - III
	423	76.71	13.23	19.5	100	Subjects included in Estimations IV - VIIb & XIV
	226	84.98	5.42	73	100	Subjects included in Estimations VIII - XIII
Invitation to Interview	504	.540	-	0	1	Full
	429	.566	-	0	1	Subjects included in Estimations I - III
	423	.572	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	1	-	1	1	Subjects included in Estimations VIII - XIII
Scholarship Awarded	504	.290	-	0	1	Full
	429	.312	-	0	1	Subjects included in Estimations I - III
	423	.314	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.593	-	0	1	Subjects included in Estimations VIII - XIII
INDEPENDENT VARIABLES						
Academic Achievement						
High School Grade Average	504	1.80	.577	1	3.7	Full
	429	1.76	.553	1	3.7	Subjects included in Estimations I - III
	423	1.76	.555	1	3.7	Subjects included in Estimations IV - VIIb & XIV
	226	1.55	.452	1	2.7	Subjects included in Estimations VIII - XIII
(Preliminary) University Grade Average	489	1.94	.551	1	3.67	Full
	429	1.89	.513	1	3.66	Subjects included in Estimations I - III
	423	1.89	.511	1	3.66	Subjects included in Estimations IV - VIIb & XIV
	226	1.68	.424	1	3.05	Subjects included in Estimations VIII - XIII
Field of Studies						
Engineering	504	.141	-	0	1	Full
	429	.140	-	0	1	Subjects included in Estimations I - III
	423	.142	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.142	-	0	1	Subjects included in Estimations VIII - XIII
Mathematics, Informatics and Natural Sciences	504	.169	-	0	1	Full
	429	.166	-	0	1	Subjects included in Estimations I - III
	423	.168	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.186	-	0	1	Subjects included in Estimations VIII - XIII
Law, Economics and Social Sciences	504	.373	-	0	1	Full
	429	.380	-	0	1	Subjects included in Estimations I - III
	423	.383	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.363	-	0	1	Subjects included in Estimations VIII - XIII
Linguistic and Cultural Sciences	504	.284	-	0	1	Full
	429	.289	-	0	1	Subjects included in Estimations I - III
	423	.284	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.288	-	0	1	Subjects included in Estimations VIII - XIII
Other	504	.034	-	0	1	Full
	429	.026	-	0	1	Subjects included in Estimations I - III
	423	.024	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.022	-	0	1	Subjects included in Estimations VIII - XIII

Type of Home Institution						
University of Applied Sciences	500	.082	-	0	1	Full
	429	.089	-	0	1	Subjects included in Estimations I - III
	423	.090	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.088	-	0	1	Subjects included in Estimations VIII - XIII
Technical University	500	.154	-	0	1	Full
	429	.154	-	0	1	Subjects included in Estimations I - III
	423	.156	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.164	-	0	1	Subjects included in Estimations VIII - XIII
University	500	.764	-	0	1	Full
	429	.758	-	0	1	Subjects included in Estimations I - III
	423	.754	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.748	-	0	1	Subjects included in Estimations VIII - XIII
Duration of Study						
Semester	504	4.39	1.69	2	14	Full
	429	4.39	1.60	2	14	Subjects included in Estimations I - III
	423	4.40	1.61	2	14	Subjects included in Estimations IV - VIIb & XIV
	226	4.46	1.46	2	10	Subjects included in Estimations VIII - XIII
Work Experience Part-Time						
(Previous or Current) Part-Time Job only at University	503	.157	-	0	1	Full
	429	.163	-	0	1	Subjects included in Estimations I - III
	423	.163	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.186	-	0	1	Subjects included in Estimations VIII - XIII
(Previous or Current) Part-Time Jobs both at & outside University	503	.193	-	0	1	Full
	429	.184	-	0	1	Subjects included in Estimations I - III
	423	.184	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.248	-	0	1	Subjects included in Estimations VIII - XIII
(Previous or Current) Part-Time Job only outside University	503	.416	-	0	1	Full
	429	.424	-	0	1	Subjects included in Estimations I - III
	423	.421	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.350	-	0	1	Subjects included in Estimations VIII - XIII
No Part-Time Job	503	.235	-	0	1	Full
	429	.228	-	0	1	Subjects included in Estimations I - III
	423	.232	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.217	-	0	1	Subjects included in Estimations VIII - XIII
Other Scholarships						
'BAföG'	504	.226	-	0	1	Full
	429	.221	-	0	1	Subjects included in Estimations I - III
	423	.222	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.190	-	0	1	Subjects included in Estimations VIII - XIII
Previous Merit-Based Stipend	504	.147	-	0	1	Full
	429	.161	-	0	1	Subjects included in Estimations I - III
	423	.163	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.243	-	0	1	Subjects included in Estimations VIII - XIII
'Studienstiftung des Dt. Volkes'	504	.060	-	0	1	Full
	429	.063	-	0	1	Subjects included in Estimations I - III
	423	.064	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.102	-	0	1	Subjects included in Estimations VIII - XIII
Also Applied for Other Scholarships	504	.421	-	0	1	Full
	429	.434	-	0	1	Subjects included in Estimations I - III
	423	.433	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.504	-	0	1	Subjects included in Estimations VIII - XIII

Number of Extracurricular Activities (ECAs)						
One Type of ECAs	504	.290	-	0	1	Full
	429	.284	-	0	1	Subjects included in Estimations I - III
	423	.286	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.270	-	0	1	Subjects included in Estimations VIII - XIII
More than one Type of ECAs	504	.516	-	0	1	Full
	429	.538	-	0	1	Subjects included in Estimations I - III
	423	.537	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.593	-	0	1	Subjects included in Estimations VIII - XIII
No Extracurricular Activity	504	.194	-	0	1	Full
	429	.177	-	0	1	Subjects included in Estimations I - III
	423	.177	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.137	-	0	1	Subjects included in Estimations VIII - XIII
Letter of Reference						
Recommendation	477	9.25	.708	5	10	Full
	429	9.26	.698	5	10	Subjects included in Estimations I - III
	423	9.27	.687	5	10	Subjects included in Estimations IV - VIIb & XIV
	226	9.47	.465	8	10	Subjects included in Estimations VIII - XIII
Recommending Person=Professor	504	.681	-	0	1	Full
	429	.697	-	0	1	Subjects included in Estimations I - III
	423	.704	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.752	-	0	1	Subjects included in Estimations VIII - XIII
Good Relation to Recommending Person	475	.834	-	0	1	Full
	429	.851	-	0	1	Subjects included in Estimations I - III
	423	.851	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.916	-	0	1	Subjects included in Estimations VIII - XIII
Language Proficiency						
Language Skills	491	88.61	9.35	40	100	Full
	429	88.85	9.21	40	100	Subjects included in Estimations I - III
	423	88.76	9.23	40	100	Subjects included in Estimations IV - VIIb & XIV
	226	90.82	7.00	60	100	Subjects included in Estimations VIII - XIII
TOEFL	504	.562	-	0	1	Full
	429	.580	-	0	1	Subjects included in Estimations I - III
	423	.582	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.770	-	0	1	Subjects included in Estimations VIII - XIII
Project-Specific Statements						
Tuition Fees (in Th. $US)	494	14.221	13.98	0	80	Full
	429	14.643	14.18	0	80	Subjects included in Estimations I - III
	423	14.652	14.18	0	80	Subjects included in Estimations IV - VIIb & XIV
	226	16.081	14.39	0	80	Subjects included in Estimations VIII - XIII
Private Guest Institution	504	.276	-	0	1	Full
	429	.289	-	0	1	Subjects included in Estimations I - III
	423	.286	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.296	-	0	1	Subjects included in Estimations VIII - XIII
Top10 Guest Institution	504	.181	-	0	1	Full
	429	.184	-	0	1	Subjects included in Estimations I - III
	423	.187	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.181	-	0	1	Subjects included in Estimations VIII - XIII

Guest Institution in Canada	504	.206	-	0	1	Full
	429	.198	-	0	1	Subjects included in Estimations I - III
	423	.196	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.190	-	0	1	Subjects included in Estimations VIII - XIII
Guest Institution in US-Midwest	504	.131	-	0	1	Full
	429	.131	-	0	1	Subjects included in Estimations I - III
	423	.130	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.119	-	0	1	Subjects included in Estimations VIII - XIII
Guest Institution in US-Northeast	504	.236	-	0	1	Full
	429	.235	-	0	1	Subjects included in Estimations I - III
	423	.235	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.265	-	0	1	Subjects included in Estimations VIII - XIII
Guest Institution in US-South	504	.177	-	0	1	Full
	429	.179	-	0	1	Subjects included in Estimations I - III
	423	.180	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.186	-	0	1	Subjects included in Estimations VIII - XIII
Guest Institution in US-West	504	.250	-	0	1	Full
	429	.256	-	0	1	Subjects included in Estimations I - III
	423	.260	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.239	-	0	1	Subjects included in Estimations VIII - XIII
Applicant Characteristics						
Gender	504	.486	-	0	1	Full
	429	.471	-	0	1	Subjects included in Estimations I - III
	423	.466	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.469	-	0	1	Subjects included in Estimations VIII - XIII
Glasses	499	.178	-	0	1	Full
	429	.186	-	0	1	Subjects included in Estimations I - III
	423	.187	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.221	-	0	1	Subjects included in Estimations VIII - XIII
Born in Germany	504	.918	-	0	1	Full
	429	.923	-	0	1	Subjects included in Estimations I - III
	423	.924	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.947	-	0	1	Subjects included in Estimations VIII - XIII
Age	504	21.87	1.45	19	32	Full
	429	21.84	1.36	19	27	Subjects included in Estimations I - III
	423	21.85	1.36	19	27	Subjects included in Estimations IV - VIIb & XIV
	226	21.69	1.19	19	27	Subjects included in Estimations VIII - XIII
Parents=Academics	504	.236	-	0	1	Full
	429	.235	-	0	1	Subjects included in Estimations I - III
	423	.227	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.199	-	0	1	Subjects included in Estimations VIII - XIII
Pre-Selection Evaluator Characteristics and Pre-Selection Situation						
Evaluator Gender	504	.188	-	0	1	Full
	429	.184	-	0	1	Subjects included in Estimations I - III
	423	.187	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.195	-	0	1	Subjects included in Estimations VIII - XIII
Evaluator Age	498	53.56	9.33	33	68	Full
	429	53.61	9.26	33	68	Subjects included in Estimations I - III
	423	53.61	9.26	33	68	Subjects included in Estimations IV - VIIb & XIV
	226	53.85	9.57	33	68	Subjects included in Estimations VIII - XIII

Evaluation Committee Characteristics						
Size of Evaluation Committee	254	4.35	0.941	3	7	Full
	226	4.36	0.939	3	7	Subjects included in Estimations VIII - XIII
Fraction of Female Evaluators	254	0.260	0.210	0	1	Full
	226	0.257	0.208	0	1	Subjects included in Estimations VIII - XIII
Mainly Male Evaluators	254	.815	-	0	1	Full
	226	.819	-	0	1	Subjects included in Estimations VIII - XIII
Average Evaluator Age	254	52.41	5.42	43.5	62	Full
	226	52.41	5.41	43.5	62	Subjects included in Estimations VIII - XIII
Dispersion Evaluator Age	254	16.02	6.25	4.24	27.48	Full
	226	15.96	6.25	4.24	27.48	Subjects included in Estimations VIII - XIII
Interview Framework						
Interview Time: 09-09:59 am	254	.134	-	0	1	Full
	226	.133	-	0	1	Subjects included in Estimations VIII - XIII
Interview Time: 10-10:59 am	254	.150	-	0	1	Full
	226	.159	-	0	1	Subjects included in Estimations VIII - XIII
Interview Time: 11-11:59 am	254	.157	-	0	1	Full
	226	.155	-	0	1	Subjects included in Estimations VIII - XIII
Interview Time: 12-01:30 pm	254	.154	-	0	1	Full
	226	.159	-	0	1	Subjects included in Estimations VIII - XIII
Interview Time: 02-02:59 pm	254	.146	-	0	1	Full
	226	.142	-	0	1	Subjects included in Estimations VIII - XIII
Interview Time: 03-03:59 pm	254	.087	-	0	1	Full
	226	.080	-	0	1	Subjects included in Estimations VIII - XIII
Interview Time: 04-04:59 pm	254	.083	-	0	1	Full
	226	.088	-	0	1	Subjects included in Estimations VIII - XIII
Interview Time: 05-06:59 pm	254	.091	-	0	1	Full
	226	.084	-	0	1	Subjects included in Estimations VIII - XIII
Year						
Selection Year	504	.518	-	0	1	Full
	429	.510	-	0	1	Subjects included in Estimations I - III
	423	.518	-	0	1	Subjects included in Estimations IV - VIIb & XIV
	226	.509	-	0	1	Subjects included in Estimations VIII - XIII

EIDESSTATTLICHE ERKLÄRUNG

Hiermit versichere ich, Linda Sophia Kurze, die vorliegende Arbeit selbstständig und unter ausschließlicher Verwendung der angegebenen Literatur und Hilfsmittel erstellt zu haben. Alle Stellen, die wörtlich oder sinngemäß veröffentlichtem oder unveröffentlichtem Schrifttum entnommen sind, habe ich als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Paderborn, der 12. Februar 2013 (Linda Sophia Kurze)